THE BULLETIN OF Mathematical BIOPHYSICS

DECEMBER 1943

THE UNIVERSITY OF CHICAGO PRESS · CHICAGO · ILLINOIS

VOLUME 5 · · · · · NUMBER 4

THE BULLETIN OF MATHEMATICAL BIOPHYSICS

EDITED BY N. RASHEVSKY

The Bulletin is devoted to publications of research in Mathematical Biophysics, as described on the inside back cover.

THE BULLETIN is published by the University of Chicago at the University of Chicago Press, 5750 Ellis Avenue, Chicago, Illinois, quarterly, in March, June, September, December. ¶The subscription price is $2.50 per year, the price of single copies is 75 cents. Orders for service of less than a full year will be charged at the single-copy rate. ¶Patrons are requested to make all remittances payable to The University of Chicago Press in postal or express money orders or bank drafts.

THE FOLLOWING are authorized agents:

For the British Empire, except North America, India, and Australasia: The Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1. Prices of yearly subscriptions and of single copies may be had on application.

CLAIMS FOR MISSING NUMBERS should be made within the month following the regular month of publication. The publishers expect to supply missing numbers free only when losses have been sustained in transit, and when the reserve stock will permit.

BUSINESS CORRESPONDENCE should be addressed to The University of Chicago Press, Chicago, Ill.

COMMUNICATIONS FOR THE EDITOR and manuscripts should be addressed to N. Rashevsky, Editorial Office of the Bulletin of Mathematical Biophysics, 5822 Drexel Avenue, Chicago, Ill.

IMPORTANT ANNOUNCEMENT

The increase in the output of papers in the field of mathematical biophysics makes it difficult to insure prompt publication without an increase in the size of the journal. Therefore, the Bulletin of Mathematical Biophysics inaugurates the following service:

Upon acceptance of a paper, the Editor, if necessary, will ask the author to shorten the paper to an extent dictated by the requirements of a reasonably prompt publication. The shortening should in no case reduce the paper to a mere abstract. Such a shortened paper will be published within six months or less.

The unabbreviated original manuscript will be kept on file at the editorial office. Any person desiring to avail himself of the complete manuscript, may obtain promptly a microfilm copy of the latter, at the cost of 1¢ per page plus postage, by applying to the Editorial Office, 5822 Drexel Avenue, Chicago, Illinois.

All papers in the Bulletin which have been thus shortened, will be marked at the end by the symbol MF, followed by a figure, indicating the number of double-spaced typewritten pages of the unabbreviated manuscript.

PRINTED BY THE DENTAN PRINTING COMPANY · · · COLORADO SPRINGS, COLORADO

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE,
DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE,
AND THE UNIVERSITY OF CHICAGO

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

I. Introduction

Theoretical neurophysiology rests on certain cardinal assumptions. The nervous system is a net of neurons, each having a soma and an axon. Their adjunctions, or synapses, are always between the axon of one neuron and the soma of another. At any instant a neuron has some threshold, which excitation must exceed to initiate an impulse. This, except for the fact and the time of its occurrence, is determined by the neuron, not by the excitation. From the point of excitation the impulse is propagated to all parts of the neuron. The velocity along the axon varies directly with its diameter, from less than one meter per second in thin axons, which are usually short, to more than 150 meters per second in thick axons, which are usually long. The time for axonal conduction is consequently of little importance in determining the time of arrival of impulses at points unequally remote from the same source. Excitation across synapses occurs predominantly from axonal terminations to somata. It is still a moot point whether this depends upon irreciprocity of individual synapses or merely upon prevalent anatomical configurations. To suppose the latter requires no hypothesis *ad hoc* and explains known exceptions, but any assumption as to cause is compatible with the calculus to come. No case is known in which excitation through a single synapse has elicited a nervous impulse in any neuron, whereas any

neuron may be excited by impulses arriving at a sufficient number of neighboring synapses within the period of latent addition, which lasts less than one quarter of a millisecond. Observed temporal summation of impulses at greater intervals is impossible for single neurons and empirically depends upon structural properties of the net. Between the arrival of impulses upon a neuron and its own propagated impulse there is a synaptic delay of more than half a millisecond. During the first part of the nervous impulse the neuron is absolutely refractory to any stimulation. Thereafter its excitability returns rapidly, in some cases reaching a value above normal from which it sinks again to a subnormal value, whence it returns slowly to normal. Frequent activity augments this subnormality. Such specificity as is possessed by nervous impulses depends solely upon their time and place and not on any other specificity of nervous energies. Of late only inhibition has been seriously adduced to contravene this thesis. Inhibition is the termination or prevention of the activity of one group of neurons by concurrent or antecedent activity of a second group. Until recently this could be explained on the supposition that previous activity of neurons of the second group might so raise the thresholds of internuncial neurons that they could no longer be excited by neurons of the first group, whereas the impulses of the first group must sum with the impulses of these internuncials to excite the now inhibited neurons. Today, some inhibitions have been shown to consume less than one millisecond. This excludes internuncials and requires synapses through which impulses inhibit that neuron which is being stimulated by impulses through other synapses. As yet experiment has not shown whether the refractoriness is relative or absolute. We will assume the latter and demonstrate that the difference is immaterial to our argument. Either variety of refractoriness can be accounted for in either of two ways. The "inhibitory synapse" may be of such a kind as to produce a substance which raises the threshold of the neuron, or it may be so placed that the local disturbance produced by its excitation opposes the alteration induced by the otherwise excitatory synapses. Inasmuch as position is already known to have such effects in the case of electrical stimulation, the first hypothesis is to be excluded unless and until it be substantiated, for the second involves no new hypothesis. We have, then, two explanations of inhibition based on the same general premises, differing only in the assumed nervous nets and, consequently, in the time required for inhibition. Hereafter we shall refer to such nervous nets as *equivalent in the extended sense.* Since we are concerned with properties of nets which are invariant under equivalence, we may make the physical assumptions which are most convenient for the calculus.

Many years ago one of us, by considerations impertinent to this argument, was led to conceive of the response of any neuron as factually equivalent to a proposition which proposed its adequate stimulus. He therefore attempted to record the behavior of complicated nets in the notation of the symbolic logic of propositions. The "all-or-none" law of nervous activity is sufficient to insure that the activity of any neuron may be represented as a proposition. Physiological relations existing among nervous activities correspond, of course, to relations among the propositions; and the utility of the representation depends upon the identity of these relations with those of the logic of propositions. To each reaction of any neuron there is a corresponding assertion of a simple proposition. This, in turn, implies either some other simple proposition or the disjunction or the conjunction, with or without negation, of similar propositions, according to the configuration of the synapses upon and the threshold of the neuron in question. Two difficulties appeared. The first concerns facilitation and extinction, in which antecedent activity temporarily alters responsiveness to subsequent stimulation of one and the same part of the net. The second concerns learning, in which activities concurrent at some previous time have altered the net permanently, so that a stimulus which would previously have been inadequate is now adequate. But for nets undergoing both alterations, we can substitute equivalent fictitious nets composed of neurons whose connections and thresholds are unaltered. But one point must be made clear: neither of us conceives the formal equivalence to be a factual explanation. *Per contra!*—we regard facilitation and extinction as dependent upon continuous changes in threshold related to electrical and chemical variables, such as after-potentials and ionic concentrations; and learning as an enduring change which can survive sleep, anaesthesia, convlusions and coma. The importance of the formal equivalence lies in this: that the alterations actually underlying facilitation, extinction and learning in no way affect the conclusions which follow from the formal treatment of the activity of nervous nets, and the relations of the corresponding propositions remain those of the logic of propositions.

The nervous system contains many circular paths, whose activity so regenerates the excitation of any participant neuron that reference to time past becomes indefinite, although it still implies that afferent activity has realized one of a certain class of configurations over time. Precise specification of these implications by means of recursive functions, and determination of those that can be embodied in the activity of nervous nets, completes the theory.

II. The Theory: Nets Without Circles

We shall make the following physical assumptions for our calculus.

1. The activity of the neuron is an "all-or-none" process.

2. A certain fixed number of synapses must be excited within the period of latent addition in order to excite a neuron at any time, and this number is independent of previous activity and position on the neuron.

3. The only significant delay within the nervous system is synaptic delay.

4. The activity of any inhibitory synapse absolutely prevents excitation of the neuron at that time.

5. The structure of the net does not change with time.

To present the theory, the most appropriate symbolism is that of Language II of R. Carnap (1938), augmented with various notations drawn from B. Russell and A. N. Whitehead (1927), including the *Principia* conventions for dots. Typographical necessity, however, will compel us to use the upright '*E*' for the existential operator instead of the inverted, and an arrow ('$\rightarrow$') for implication instead of the horseshoe. We shall also use the Carnap syntactical notations, but print them in boldface rather than German type; and we shall introduce a functor S, whose value for a property P is the property which holds of a number when P holds of its predecessor; it is defined by '$S(P)\,(t) \mathrel{.\equiv.} P(Kx)\ .\ t = x'$)'; the brackets around its argument will often be omitted, in which case this is understood to be the nearest predicate-expression [***Pr***] on the right. Moreover, we shall write $S^2\boldsymbol{Pr}$ for $S(S(\boldsymbol{Pr}))$, etc.

The neurons of a given net $\mathcal{N}$ may be assigned designations 'c_1', 'c_2', $\cdots$, 'c_n'. This done, we shall denote the property of a number, that a neuron c_i fires at a time which is that number of synaptic delays from the origin of time, by '*N*' with the numeral i as subscript, so that $N_i(t)$ asserts that c_i fires at the time t. N_i is called the *action* of c_i. We shall sometimes regard the subscripted numeral of '*N*' as if it belonged to the object-language, and were in a place for a functoral argument, so that it might be replaced by a number-variable [*z*] and quantified; this enables us to abbreviate long but finite disjunctions and conjunctions by the use of an operator. We shall employ this locution quite generally for sequences of ***Pr***; it may be secured formally by an obvious disjunctive definition. The predicates 'N_1', 'N_2', $\cdots$, comprise the syntactical class '*N*'.

Let us define the *peripheral afferents* of $\mathcal{N}$ as the neurons of $\mathcal{N}$ with no axons synapsing upon them. Let $N_1, \cdots, N_p$ denote the actions of such neurons and $N_{p+1}, N_{p+2}, \cdots, N_n$ those of the rest. Then a *solution of* $\mathcal{N}$ will be a class of sentences of the form S_i: $N_{p+1}(z_1) .\equiv. Pr_i(N_1, N_2, \cdots, N_p, z_1)$, where Pr_i contains no free variable save z_1 and no descriptive symbols save the N in the argument [Arg], and possibly some constant sentences [sa]; and such that each S_i is true of $\mathcal{N}$. Conversely, given a $Pr_1({}^1p^1_1, {}^1p^1_2, \cdots, {}^1p^1_p, z_1, s)$, containing no free variable save those in its Arg, we shall say that it is *realizable in the narrow sense* if there exists a net $\mathcal{N}$ and a series of N_i in it such that $N_1(z_1) .\equiv. Pr_1(N_1, N_2, \cdots, z_1, sa_1)$ is true of it, where sa_1 has the form $N(0)$. We shall call it *realizable in the extended sense,* or simply *realizable,* if for some n $S^n(Pr_1)(p_1, \cdots, p_p, z_1, s)$ is realizable in the above sense. c_{pi} is here the realizing neuron. We shall say of two laws of nervous excitation which are such that every S which is realizable in either sense upon one supposition is also realizable, perhaps by a different net, upon the other, that they are equivalent assumptions, in that sense.

The following theorems about realizability all refer to the extended sense. In some cases, sharper theorems about narrow realizability can be obtained; but in addition to greater complication in statement this were of little practical value, since our present neurophysiological knowledge determines the law of excitation only to extended equivalence, and the more precise theorems differ according to which possible assumption we make. Our less precise theorems, however, are invariant under equivalence, and are still sufficient for all purposes in which the exact time for impulses to pass through the whole net is not crucial.

Our central problems may now be stated exactly: first, to find an effective method of obtaining a set of computable S constituting a solution of any given net; and second, to characterize the class of realizable S in an effective fashion. Materially stated, the problems are to calculate the behavior of any net, and to find a net which will behave in a specified way, when such a net exists.

A net will be called *cyclic* if it contains a circle: i.e., if there exists a chain $c_i, c_{i+1}, \cdots$ of neurons on it, each member of the chain synapsing upon the next, with the same beginning and end. If a set of its neurons $c_1, c_2, \cdots, c_p$ is such that its removal from $\mathcal{N}$ leaves it without circles, and no smaller class of neurons has this property, the set is called a *cyclic* set, and its cardinality is the *order* of $\mathcal{N}$. In an important sense, as we shall see, the order of a net is an index of the complexity of its behavior. In particular, nets of zero order have especially simple properties; we shall discuss them first.

Let us define a *temporal propositional expression* (a *TPE*), designating a *temporal propositional function* (*TPF*), by the following recursion:

1. A ${}^1p^1[z_1]$ is a *TPE*, where p_1 is a predicate-variable.
2. If S_1 and S_2 are *TPE* containing the same free individual variable, so are SS_1, $S_1 \mathrm{v} S_2$, $S_1.S_2$ and $S_1 . \circ S_2$.
3. Nothing else is a *TPE*.

THEOREM I.

Every net of order 0 can be solved in terms of temporal propositional expressions.

Let c_i be any neuron of $\mathcal{N}$ with a threshold $\theta_i > 0$, and let c_{i1}, c_{i2}, $\cdots$, c_{ip} have respectively n_{i1}, n_{i2}, $\cdots$, n_{ip} excitatory synapses upon it. Let c_{j1}, c_{j2}, $\cdots$, c_{jq} have inhibitory synapses upon it. Let κ_i be the set of the subclasses of $\{n_{i1}, n_{i2}, \cdots, n_{ip}\}$ such that the sum of their members exceeds θ_i. We shall then be able to write, in accordance with the assumptions mentioned above,

$$N_i(z_1) . \equiv . S\left\{\prod_{m=1}^{q} \circ N_{jm}(z_1) . \sum_{a \varepsilon K_i} \prod_{s \varepsilon a} N_{is}(z_1)\right\} \tag{1}$$

where the 'Σ' and 'Π' are syntactical symbols for disjunctions and conjunctions which are finite in each case. Since an expression of this form can be written for each c_i which is not a peripheral afferent, we can, by substituting the corresponding expression in (1) for each N_{jm} or N_{is} whose neuron is not a peripheral afferent, and repeating the process on the result, ultimately come to an expression for N_i in terms solely of peripherally afferent N, since $\mathcal{N}$ is without circles. Moreover, this expression will be a *TPE*, since obviously (1) is; and it follows immediately from the definition that the result of substituting a *TPE* for a constituent $p(z)$ in a *TPE* is also one.

THEOREM II.

Every TPE *is realizable by a net of order zero.*

The functor S obviously commutes with disjunction, conjunction, and negation. It is obvious that the result of substituting any S_i, realizable in the narrow sense (i.n.s.), for the $p(z)$ in a realizable expression S_1 is itself realizable i.n.s.; one constructs the realizing net by replacing the peripheral afferents in the net for S_1 by the realizing neurons in the nets for the S_i. The one neuron net realizes $p_1(z_1)$ i.n.s.,

and Figure 1-a shows a net that realizes $Sp_1(z_1)$ and hence SS_2, i.n.s., if S_2 can be realized i.n.s. Now if S_2 and S_3 are realizable then S^mS_2 and S^nS_3 are realizable i.n.s., for suitable m and n. Hence so are $S^{m+n}S_2$ and $S^{m+n}S_3$. Now the nets of Figures 1b, c and d respectively realize $S(p_1(z_1) \vee p_2(z_1))$, $S(p_1(z_1) . p_2(z_1))$, and $S(p_1(z_1) . \sim p_2(z_1))$ i.n.s. Hence $S^{m+n+1}(S_1 \vee S_2)$, $S^{m+n+1}(S_1 . S_2)$, and $S^{m+n+1}(S_1 . \sim S_2)$ are realizable i.n.s. Therefore $S_1 \vee S_2\ S_1 . S_2\ S_1 . \sim S_2$ are realizable if S_1 and S_2 are. By complete induction, all *TPE* are realizable. In this way all nets may be regarded as built out of the fundamental elements of Figures 1a, b, c, d, precisely as the temporal propositional expressions are generated out of the operations of precession, disjunction, conjunction, and conjoined negation. In particular, corresponding to any description of state, or distribution of the values *true* and *false* for the actions of all the neurons of a net save that which makes them all false, a single neuron is constructible whose firing is a necessary and sufficient condition for the validity of that description. Moreover, there is always an indefinite number of topologically different nets realizing any *TPE*.

THEOREM III.

Let there be given a complex sentence S_1 *built up in any manner out of elementary sentences of the form* $\mathbf{p}(\mathbf{z}_1 - \mathbf{zz})$ *where* **zz** *is any numeral, by any of the propositional connections: negation, disjunction, conjunction, implication, and equivalence. Then* S_1 *is a* TPE *and only if it is false when its constituent* $\mathbf{p}(\mathbf{z}_1 - \mathbf{zz})$ *are all assumed false—i.e., replaced by false sentences — or that the last line in its truth-table contains an* '**F**',*—or there is no term in its Hilbert disjunctive normal form composed exclusively of negated terms.*

These latter three conditions are of course equivalent (Hilbert and Ackermann, 1938). We see by induction that the first of them is necessary, since $p(z_1 - zz)$ becomes false when it is replaced by a false sentence, and $S_1 \vee S_2$, $S_1 . S_2$ and $S_1 . \sim S_2$ are all false if both their constituents are. We see that the last condition is sufficient by remarking that a disjunction is a *TPE* when its constituents are, and that any term

$$S_1 . S_2 . \ldots . S_m . \sim S_{m+1} . \sim \ldots . \sim S_n$$

can be written as

$$(S_1 . S_2 . \ldots . S_m) . \sim (S_{m+1} \vee S_{m+2} \vee \ldots \vee S_n),$$

which is clearly a *TPE*.

The method of the last theorems does in fact provide a very convenient and workable procedure for constructing nervous nets to or-

der, for those cases where there is no reference to events indefinitely far in the past in the specification of the conditions. By way of example, we may consider the case of heat produced by a transient cooling.

If a cold object is held to the skin for a moment and removed, a sensation of heat will be felt; if it is applied for a longer time, the sensation will be only of cold, with no preliminary warmth, however transient. It is known that one cutaneous receptor is affected by heat, and another by cold. If we let N_1 and N_2 be the actions of the respective receptors and N_3 and N_4 of neurons whose activity implies a sensation of heat and cold, our requirements may be written as

$$N_3(t) \,:\equiv:\, N_1(t-1) \,.\, \mathrm{v} \,.\, N_2(t-3) \,.\, \sim N_2(t-2)$$
$$N_4(t) \,.\equiv.\, N_2(t-2) \,.\, N_2(t-1)$$

where we suppose for simplicity that the required persistence in the sensation of cold is say two synaptic delays, compared with one for that of heat. These conditions clearly fall under Theorem III. A net may consequently be constructed to realize them, by the method of Theorem II. We begin by writing them in a fashion which exhibits them as built out of their constituents by the operations realized in Figures 1a, b, c, d: i.e., in the form

$$N_3(t) \,.\equiv.\, S\{N_1(t) \,\mathrm{v}\, S[(SN_2(t)) \,.\, \sim N_2(t)]\}$$
$$N_4(t) \,.\equiv.\, S\{[SN_2(t)] \,.\, N_2(t)\}\,.$$

First we construct a net for the function enclosed in the greatest number of brackets and proceed outward; in this case we run a net of the form shown in Figure 1a from c_2 to some neuron c_a, say, so that

$$N_a(t) \,.\equiv.\, SN_2(t).$$

Next introduce two nets of the forms 1c and 1d, both running from c_a and c_2, and ending respectively at c_4 and say c_b. Then

$$N_4(t) \,.\equiv.\, S[N_a(t) \,.\, N_2(t)] \,.\equiv.\, S[(SN_2(t)) \,.\, N_2(t)].$$
$$N_b(t) \,.\equiv.\, S[N_a(t) \,.\, \sim N_2(t)] \,.\equiv.\, S[(SN_2(t)) \,.\, \sim N_2(t)].$$

Finally, run a net of the form 1b from c_1 and c_b to c_3, and derive

$$N_3(t) \,.\equiv.\, S[N_1(t) \,\mathrm{v}\, N_b(t)]$$
$$.\equiv.\, S\{N_1(t) \,\mathrm{v}\, S[(SN_2(t)] \,.\, \sim N_2(t)\}\,.$$

These expressions for $N_3(t)$ and $N_4(t)$ are the ones desired; and the realizing net *in toto* is shown in Figure 1e.

This illusion makes very clear the dependence of the correspondence between perception and the "external world" upon the specific structural properties of the intervening nervous net. The same illu-

sion, of course, could also have been produced under various other assumptions about the behavior of the cutaneous receptors, with correspondingly different nets.

We shall now consider some theorems of equivalence: i.e., theorems which demonstrate the essential identity, save for time, of various alternative laws of nervous excitation. Let us first discuss the case of *relative inhibition.* By this we mean the supposition that the firing of an inhibitory synapse does not absolutely prevent the firing of the neuron, but merely raises its threshold, so that a greater number of excitatory synapses must fire concurrently to fire it than would otherwise be needed. We may suppose, losing no generality, that the increase in threshold is unity for the firing of each such synapse; we then have the theorem:

THEOREM IV.

Relative and absolute inhibition are equivalent in the extended sense.

We may write out a law of nervous excitation after the fashion of (1), but employing the assumption of relative inhibition instead; inspection then shows that this expression is a *TPE.* An example of the replacement of relative inhibition by absolute is given by Figure 1f. The reverse replacement is even easier; we give the inhibitory axons afferent to c_i any sufficiently large number of inhibitory synapses apiece.

Second, we consider the case of extinction. We may write this in the form of a variation in the threshold θ_i; after the neuron c_i has fired; to the nearest integer—and only to this approximation is the variation in threshold significant in natural forms of excitation—this may be written as a sequence $\theta_i + b_j$ for j synaptic delays after firing, where $b_j = 0$ for j large enough, say $j = M$ or greater. We may then state

THEOREM V.

Extinction is equivalent to absolute inhibition.

For, assuming relative inhibition to hold for the moment, we need merely run M circuits $\mathcal{T}_1, \mathcal{T}_2, \cdots \mathcal{T}_M$ containing respectively $1, 2, \cdots, M$ neurons, such that the firing of each link in any is sufficient to fire the next, from the neuron c_i back to it, where the end of the circuit $\mathcal{T}_j$ has just b_j inhibitory synapses upon c_i. It is evident that this will produce the desired results. The reverse substitution may be accomplished by the diagram of Figure 1g. From the transitivity of

replacement, we infer the theorem. To this group of theorems also belongs the well-known

THEOREM VI.

Facilitation and temporal summation may be replaced by spatial summation.

This is obvious: one need merely introduce a suitable sequence of delaying chains, of increasing numbers of synapses, between the exciting cell and the neuron whereon temporal summation is desired to hold. The assumption of spatial summation will then give the required results. See e.g. Figure 1h. This procedure had application in showing that the observed temporal summation in gross nets does not imply such a mechanism in the interaction of individual neurons.

The phenomena of learning, which are of a character persisting over most physiological changes in nervous activity, seem to require the possibility of permanent alterations in the structure of nets. The simplest such alteration is the formation of new synapses or equivalent local depressions of threshold. We suppose that some axonal terminations cannot at first excite the succeeding neuron; but if at any time the neuron fires, and the axonal terminations are simultaneously excited, they become synapses of the ordinary kind, henceforth capable of exciting the neuron. The loss of an inhibitory synapse gives an entirely equivalent result. We shall then have

THEOREM VII.

Alterable synapses can be replaced by circles.

This is accomplished by the method of Figure 1i. It is also to be remarked that a neuron which becomes and remains spontaneously active can likewise be replaced by a circle, which is set into activity by a peripheral afferent when the activity commences, and inhibited by one when it ceases.

III. The Theory: Nets with Circles.

The treatment of nets which do not satisfy our previous assumption of freedom from circles is very much more difficult than that case. This is largely a consequence of the possibility that activity may be set up in a circuit and continue reverberating around it for an indefinite period of time, so that the realizable $\boldsymbol{Pr}$ may involve reference to past events of an indefinite degree of remoteness. Consider such a net $\mathcal{N}$, say of order p, and let $c_1, c_2, \cdots, c_p$ be a cyclic set of neurons of $\mathcal{N}$. It is first of all clear from the definition that every N_s

of $\mathcal{N}$ can be expressed as a *TPE*, of $N_1, N_2, \cdots, N_p$ and the absolute afferents; the solution of $\mathcal{N}$ involves then only the determination of expressions for the cyclic set. This done, we shall derive a set of expressions $[A]$:

$$N_i(z_1) \,.\equiv.\, Pr_i[S^{n_{i1}}N_1(z_1), S^{n_{i2}}N_2(z_1), \cdots, S^{n_{ip}}N_p(z_1)], \quad (2)$$

where Pr_i also involves peripheral afferents. Now if n is the least common multiple of the n_{ij}, we shall, by substituting their equivalents according to (2) in (3) for the N_j, and repeating this process often enough on the result, obtain S of the form

$$N_i(z_1) \,.\equiv.\, Pr_1[S^n N_1(z_1), S^n N_2(z_1), \cdots, S^n N_p(z_1)]\,. \quad (3)$$

These expressions may be written in the Hilbert disjunctive normal form as

$$N_i(z_1) \,.\equiv.\, \sum_{\substack{\alpha\varepsilon K \\ \beta_\alpha\varepsilon K}} S_\alpha \prod_{j\varepsilon K} S^n N_j(z_1) \prod_{j\varepsilon\beta_\alpha} \sim S^n N_j(z_1), \text{ for suitable } \kappa, \quad (4)$$

where S_α is a *TPE* of the absolute afferents of $\mathcal{N}$. There exist some 2^p different sentences formed out of the p N_i by conjoining to the conjunction of some set of them the conjunction of the negations of the rest. Denumerating these by $X_1(z_1), X_2(z_1), \cdots, X_{2^p}(z_1)$, we may, by use of the expressions (4), arrive at an equipollent set of equations of the form

$$X_i(z_1) \,.\equiv.\, \sum_{j=1}^{2^p} Pr_{ij}(z_1)\,.\,S^n X_j(z_1)\,. \quad (5)$$

Now we import the subscripted numerals i,j into the object-language: i.e., define Pr_1 and Pr_2 such that $Pr_1(zz_1,z_1) \,.\equiv.\, X_i(z_1)$ and $Pr_2(zz_1, zz_2, z_1) \,.\equiv.\, Pr_{ij}(z_1)$ are provable whenever zz_1 and zz_2 denote i and j respectively.

Then we may rewrite (5) as

$$\begin{aligned}&(z_1)zz_p : Pr_1(z_1, z_3)\\ &\quad .\equiv.\, (Ez_2)zz_p\,.\,Pr_2(z_1, z_2, z_3 - zz_n)\,.\,Pr_1(z_2, z_3 - zz_n)\end{aligned} \quad (6)$$

where zz_n denotes n and zz_p denotes 2^p. By repeated substitution we arrive at an expression

$$\begin{aligned}&(z_1)zz_p : Pr_1(z_1, zz_n\, zz_2) \,.\equiv.\, (Ez_2)zz_p\,(Ez_3)zz_p \ldots (Ez_n)zz_p\,.\\ &Pr_2(z_1, z_2, zz_n(zz_2 - 1))\,.\,Pr_2(z_2, z_3, zz_n(zz_2 - 1)) \ldots\ldots\end{aligned} \quad (7)$$

$Pr_2(z_{n-1}, z_n, 0)\,.\,Pr_1(z_n, 0)$, for any numeral zz_2 which denotes s. This is easily shown by induction to be equipollent to

$$(z_1)zz_p : . \boldsymbol{Pr}_1(z_1, zz_n zz_2) : \equiv : (Ef)(z_2)zz_2 - 1 f(z_2 zz_n) \leqq zz_p . f(zz_n zz_2) = z_1 . \boldsymbol{Pr}_2(f(zz_n (z_2 + 1))), f(zz_n z_2)) . \boldsymbol{Pr}_1(f(0), 0) \tag{8}$$

and since this is the case for all zz_2, it is also true that

$$(z_4)(z_1)zz_p : \boldsymbol{Pr}_1(z_1, z_4) . \equiv . (Ef)(z_2)(z_4 - 1) . f(z_2) \leqq zz_p . f(z_4) = z_1 f(z_4) = z_1 . \boldsymbol{Pr}_2[f(z_2 + 1), f(z_2), z_2]. \boldsymbol{Pr}_1[f(\mathrm{res}(z_4, zz_n)), \mathrm{res}(z_4, zz_n)], \tag{9}$$

where zz_n denotes n, $\mathrm{res}(r,s)$ is the residue of r mod s and zz_p denotes 2^p. This may be written in a less exact way as

$$N_i(t) . \equiv . (E\phi)(x)t - 1 . \phi(x) \leqq 2^p . \phi(t) = i . P[\phi(x + 1), \phi(x) . N_{\phi(0)}(0)],$$

where x and t are also assumed divisible by n, and $\boldsymbol{Pr}_2$ denotes P. From the preceding remarks we shall have

THEOREM VIII.

The expression (9) *for neurons of the cyclic set of a net* $\mathcal{N}$ *together with certain* TPE *expressing the actions of other neurons in terms of them, constitute a solution of* $\mathcal{N}$.

Consider now the question of the realizability of a set of $\boldsymbol{S}_i$. A first necessary condition, demonstrable by an easy induction, is that

$$(z_2)z_1 . \boldsymbol{p}_1(z_2) \equiv \boldsymbol{p}_2(z_2) . \rightarrow . \boldsymbol{S}_i \equiv \boldsymbol{S}_i \left\{ \begin{matrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \end{matrix} \right\} \tag{10}$$

should be true, with similar statements for the other free $\boldsymbol{p}$ in $\boldsymbol{S}_i$: i.e., no nervous net can take account of future peripheral afferents. Any $\boldsymbol{S}_i$ satisfying this requirement can be replaced by an equipollent $\boldsymbol{S}$ of the form

$$(Ef)(z_2)z_1(z_3)zz_p : f \varepsilon \boldsymbol{Pr}_{mi} : f(z_1, z_2, z_3) = 1 . \equiv . \boldsymbol{p}_{z_3}(z_2) \tag{11}$$

where zz_p denotes p, by defining

$$\boldsymbol{Pr}_{mi} = \overset{\Lambda}{f}[(z_1)(z_2)z_1(z_3)zz_p : . f(z_1, z_2, z_3) = 0 . \mathrm{v} . f(z_1, z_2, z_3) = 1 : f(z_1, z_2, z_3) = 1 . \equiv . \boldsymbol{p}_{z_3}(z_2) : \rightarrow : \boldsymbol{S}_i].$$

Consider now these series of classes α_i, for which

$$N_i(t) : \equiv : (E\phi)(x)t(m)q : \phi \varepsilon \alpha_i : N_m(x) . \equiv . \phi(t, x, m) = 1 . \quad [i = q + 1, \cdots, M] \tag{12}$$

holds for some net. These will be called *prehensible* classes. Let us define the *Boolean ring* generated by a class of classes κ as the aggregate of the classes which can be formed from members of κ by repeated application of the logical operations; i.e., we put

$$\mathcal{R}(\kappa) = p\hat{\lambda}[(\alpha,\beta):\alpha\,\varepsilon\,\kappa \rightarrow \alpha\,\varepsilon\,\lambda:\alpha,\beta\,\varepsilon\,\lambda.\rightarrow.-\alpha,\alpha.\beta,\alpha\vee\beta\,\varepsilon\,\lambda].$$

We shall also define

$$\begin{aligned}\bar{\mathcal{R}}(\kappa) &.=. \mathcal{R}(\kappa) - \iota'p' - ''\kappa,\\ \mathcal{R}_e(\kappa) &= p\,'\hat{\lambda}[(\alpha,\beta):\alpha\,\varepsilon\,\kappa\rightarrow\alpha\,\varepsilon\,\lambda.\rightarrow.-\alpha,\alpha.\beta,\alpha\vee\beta,S''\alpha\,\varepsilon\,\lambda]\\ \bar{\mathcal{R}}_e(\kappa) &= \mathcal{R}_e(\kappa) - \iota'p' - ''\kappa,\end{aligned}$$

and

$$\sigma(\psi,t) = \hat{\phi}[(m).\phi(t+1,t,m) = \psi(m)].$$

The class $\mathcal{R}_e(\kappa)$ is formed from κ in analogy with $\mathcal{R}(\kappa)$, but by repeated application not only of the logical operations but also of that which replaces a class of properties $P\,\varepsilon\,\alpha$ by $S(P)\,\varepsilon\,S''\alpha$. We shall then have the

LEMMA

$\boldsymbol{Pr}_1(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m,\boldsymbol{z}_1)$ is a *TPE* if and only if

$$\begin{aligned}(\boldsymbol{z}_1)(\boldsymbol{p}_1,\cdots,\boldsymbol{p}_m)(E\boldsymbol{p}_{m+1}) &: \boldsymbol{p}_{m+1}\,\varepsilon\,\bar{\mathcal{R}}_e(\{\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m\})\\ &\boldsymbol{p}_{m+1}(\boldsymbol{z}_1) \equiv \boldsymbol{Pr}_1(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_m,\boldsymbol{z}_1)\end{aligned} \tag{13}$$

is true; and it is a *TPE* not involving 'S' if and only if this holds when '$\bar{\mathcal{R}}_e$' is replaced by '$\bar{\mathcal{R}}$', and we then obtain

THEOREM IX.

A series of classes $\alpha_1,\alpha_2,\cdots\alpha_s$ is a series of prehensible classes if and only if

$$\begin{aligned}&(Em)(En)(p)n(i)(\psi) :. (x)m\,\psi(x) = 0\vee\psi(x) = 1 :\rightarrow: (E\beta)\\ &(Ey)m.\psi(y) = 0.\beta\,\varepsilon\,\mathcal{R}[\hat{\gamma}((Ei).\gamma = \alpha_i)).\mathrm{v}.(x)m.\\ &\psi(x) = 0.\beta\,\varepsilon\,\bar{\mathcal{R}}[\hat{\gamma}((Ei).\gamma = \alpha_i)] : (t)(\phi) : \phi\,\varepsilon\,\alpha_i.\\ &\qquad \sigma(\phi, nt+p).\rightarrow.(Ef).f\,\varepsilon\,\beta.(w)m(x)t-1.\\ &\qquad\qquad \phi(n(t+1)+p, nx+p, w) = f(nt+p, nx+p, w).\end{aligned} \tag{14}$$

The proof here follows directly from the lemma. The condition is necessary, since every net for which an expression of the form (4) can be written obviously verifies it, the ψ's being the characteristic functions of the S_α and the β for each ψ being the class whose designation has the form $\prod_{i\varepsilon\alpha} \boldsymbol{Pr}_i \prod_{j\varepsilon\beta_\alpha} \boldsymbol{Pr}_j$, where $\boldsymbol{Pr}_k$ denotes α_k for all k. Conversely, we may write an expression of the form (4) for a net $\mathcal{N}$ fulfilling prehensible classes satisfying (14) by putting for the $\boldsymbol{Pr}_\alpha$ $\boldsymbol{Pr}$ denoting the ψ's, and a $\boldsymbol{Pr}$, written in the analogue for classes of the disjunctive normal form, and denoting the α corresponding to that ψ, conjoined to it. Since every $\boldsymbol{S}$ of the form (4) is clearly realizable, we have the theorem.

It is of some interest to consider the extent to which we can by knowledge of the present determine the whole past of various special nets: i.e., when we may construct a net the firing of the cyclic set of whose neurons requires the peripheral afferents to have had a set of past values specified by given functions ϕ_i. In this case the classes α_i of the last theorem reduced to unit classes; and the condition may be transformed into

$$\begin{aligned}&(E\,m,n)\,(p)n(i,\psi)\,(Ej) : .\ (x)m : \psi(x) = 0\ .\mathbf{v}.\ \psi(x) = 1 :\\ &\quad \phi_i\,\varepsilon\,\sigma(\psi, n\,t+p) : \rightarrow : (w)m(x)t-1\ .\ \phi_i(n(t+1)\\ &\quad +p, n\,x+p, w) = \phi_j(n\,t+p, n\,x+p, w) : .\\ &\quad (u,v)\,(w)m\ .\ \phi_i(n(u+1)+p, n\,u+p, w)\\ &\quad = \phi_i(n(v+1)+p, n\,v+p, w).\end{aligned}$$

On account of limitations of space, we have presented the above argument very sketchily; we propose to expand it and certain of its implications in a further publication.

The condition of the last theorem is fairly simple in principle, though not in detail; its application to practical cases would, however, require the exploration of some 2^{2^n} classes of functions, namely the members of $\mathcal{R}(\{\alpha_1, \cdots, \alpha_s\})$. Since each of these is a possible β of Theorem IX, this result cannot be sharpened. But we may obtain a sufficient condition for the realizability of an $\boldsymbol{S}$ which is very easily applicable and probably covers most practical purposes. This is given by

THEOREM *X*.

Let us define a set $\boldsymbol{K}$ of $\boldsymbol{S}$ by the following recursion:

1. Any *TPE* and any *TPE* whose arguments have been replaced by members of $\boldsymbol{K}$ belong to $\boldsymbol{K}$;

2. If $\boldsymbol{Pr}_1(z_1)$ is a member of $\boldsymbol{K}$, then $(z_2)z_1\ .\ \boldsymbol{Pr}_1(z_2)$, $(Ez_2)z_1\ .$

$Pr_1(z_2)$, and $C_{mn}(z_1)$. s belong to it, where C_{mn} denotes the property of being congruent to m modulo n, $m < n$.

3. *The set* K *has no further members*

Then every member of K is realizable.

For, if $Pr_1(z_1)$ is realizable, nervous nets for which

$$N_i(z_1) \ . \equiv . \ Pr_1(z_1) \ . SN_i(z_1)$$
$$N_i(z_1) \ . \equiv . \ Pr_1(z_1) \ \text{v} \ SN_i(z_1)$$

are the expressions of equation (4), realize $(z_2)z_1$. $Pr_1(z_2)$ and $(E\ z_2)z_1$. $Pr_1(z_2)$ respectively; and a simple circuit, c_1, c_2, $\cdots$, c_n, of n links, each sufficient to excite the next, gives an expression

$$N_m(z_1) \ . \equiv . \ N_1(0) \ . C_{mn}$$

for the last form. By induction we derive the theorem.

One more thing is to be remarked in conclusion. It is easily shown: first, that every net, if furnished with a tape, scanners connected to afferents, and suitable efferents to perform the necessary motor-operations, can compute only such numbers as can a Turing machine; second, that each of the latter numbers can be computed by such a net; and that nets with circles can be computed by such a net; and that nets with circles can compute, without scanners and a tape, some of the numbers the machine can, but no others, and not all of them. This is of interest as affording a psychological justification of the Turing definition of computability and its equivalents, Church's λ — definability and Kleene's primitive recursiveness: If any number can be computed by an organism, it is computable by these definitions, and conversely.

IV. Consequences

Causality, which requires description of states and a law of necessary connection relating them, has appeared in several forms in several sciences, but never, except in statistics, has it been as irreciprocal as in this theory. Specification for any one time of afferent stimulation and of the activity of all constituent neurons, each an "all-or-none" affair, determines the state. Specification of the nervous net provides the law of necessary connection whereby one can compute from the description of any state that of the succeeding state, but the inclusion of disjunctive relations prevents complete determination of the one before. Moreover, the regenerative activity of constituent circles renders reference indefinite as to time past. Thus our knowledge of the world, including ourselves, is incomplete

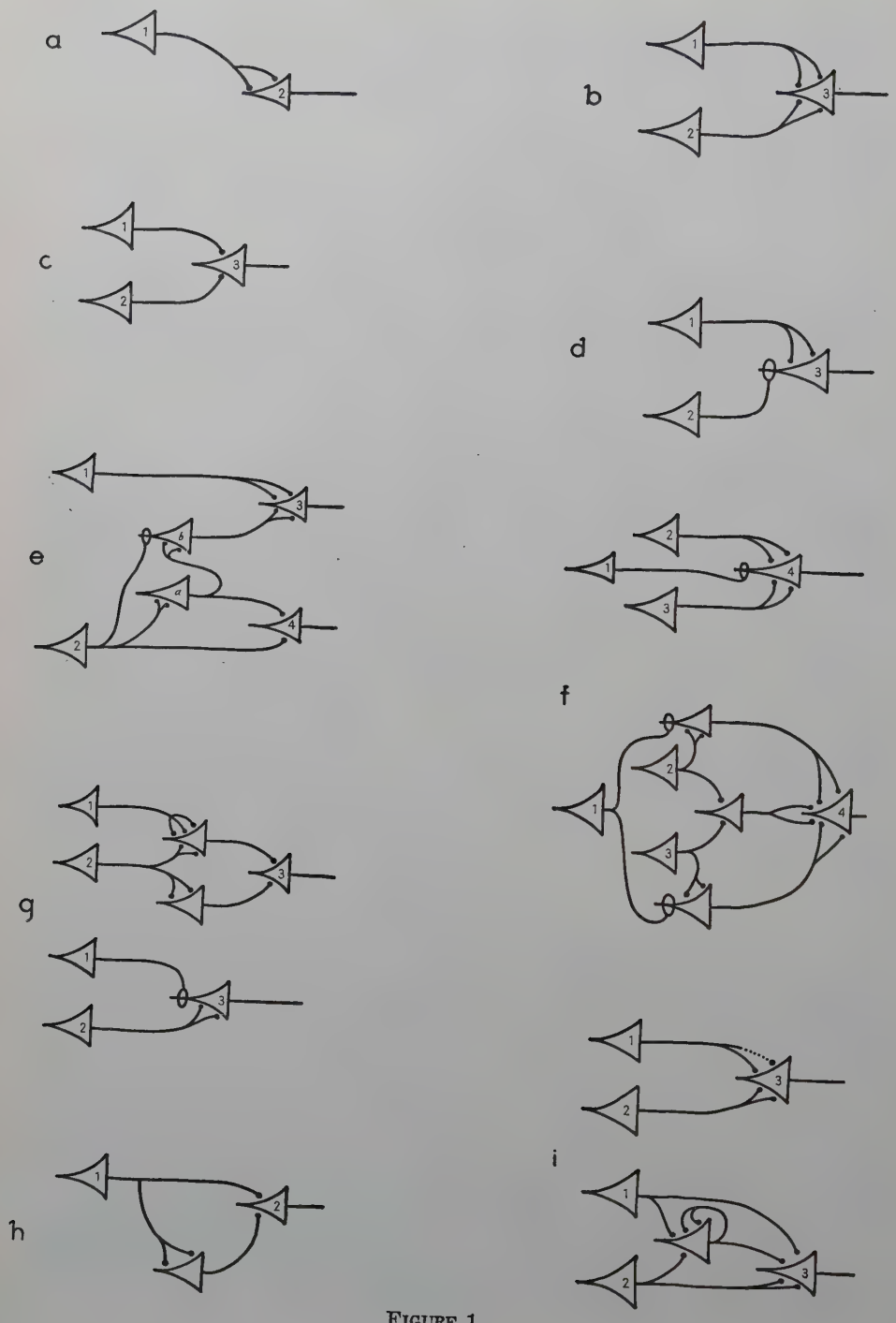

FIGURE 1

as to space and indefinite as to time. This ignorance, implicit in all our brains, is the counterpart of the abstraction which renders our knowledge useful. The role of brains in determining the epistemic relations of our theories to our observations and of these to the facts is all too clear, for it is apparent that every idea and every sensation is realized by activity within that net, and by no such activity are the actual afferents fully determined.

There is no theory we may hold and no observation we can make that will retain so much as its old defective reference to the facts if the net be altered. Tinitus, paraesthesias, hallucinations, delusions, confusions and disorientations intervene. Thus empiry confirms that if our nets are undefined, our facts are undefined, and to the "real" we can attribute not so much as one quality or "form." With determination of the net, the unknowable object of knowledge, the "thing in itself," ceases to be unknowable.

To psychology, however defined, specification of the net would contribute all that could be achieved in that field—even if the analysis were pushed to ultimate psychic units or "psychons," for a psychon can be no less than the activity of a single neuron. Since that activity is inherently propositional, all psychic events have an intentional, or "semiotic," character. The "all-or-none" law of these activities, and the conformity of their relations to those of the logic of propositions, insure that the relations of psychons are those of the two-valued logic of propositions. Thus in psychology, introspective, behavioristic or physiological, the fundamental relations are those of two-valued logic.

EXPRESSION FOR THE FIGURES

In the figure the neuron c_i is always marked with the numeral i upon the body of the cell, and the corresponding action is denoted by 'N' with i as subscript, as in the text.

Figure 1a $N_2(t) \,.\equiv.\, N_1(t-1)$

Figure 1b $N_3(t) \,.\equiv.\, N_1(t-1) \vee N_2(t-1)$

Figure 1c $N_3(t) \,.\equiv.\, N_1(t-1).N_2(t-1)$

Figure 1d $N_3(t) \,.\equiv.\, N_1(t-1).\sim N_2(t-1)$

Figure 1e $N_3(t) :\equiv: N_1(t-1) \,.\vee.\, N_2(t-3).\sim N_2(t-2)$

$N_4(t) \,.\equiv.\, N_2(t-2).N_2(t-1)$

Figure 1f $N_4(t) :\equiv: \sim N_1(t-1).\, N_2(t-1) \vee N_3(t-1) \,.\vee.\, N_1(t-1).\, N_2(t-1).\, N_3(t-1)$

$N_4(t) :\equiv: \sim N_1(t-2).\, N_2(t-2) \vee N_3(t-2) \,.\vee.\, N_1(t-2).\, N_2(t-2).\, N_3(t-2)$

Figure 1g $N_3(t) \,.\equiv.\, N_2(t-2).\sim N_1(t-3)$

Figure 1h $N_2(t) \,.\equiv.\, N_1(t-1).N_1(t-2)$

Figure 1i $N_3(t) :\equiv: N_2(t-1) \,.\vee.\, N_1(t-1).(Ex)\, t-1 \,.\, N_1(x) \,.\, N_2(x)$

Hence arise constructional solutions of holistic problems involving the differentiated continuum of sense awareness and the normative, perfective and resolvent properties of perception and execution. From the irreciprocity of causality it follows that even if the net be known, though we may predict future from present activities, we can deduce neither afferent from central, nor central from efferent, nor past from present activities—conclusions which are reinforced by the contradictory testimony of eye-witnesses, by the difficulty of diagnosing differentially the organically diseased, the hysteric and the malingerer, and by comparing one's own memories or recollections with his contemporaneous records. Moreover, systems which so respond to the difference between afferents to a regenerative net and certain activity within that net, as to reduce the difference, exhibit purposive behavior; and organisms are known to possess many such systems, subserving homeostasis, appetition and attention. Thus both the formal and the final aspects of that activity which we are wont to call *mental* are rigorously deduceable from present neurophysiology. The psychiatrist may take comfort from the obvious conclusion concerning causality—that, for prognosis, history is never necessary. He can take little from the equally valid conclusion that his observables are explicable only in terms of nervous activities which, until recently, have been beyond his ken. The crux of this ignorance is that inference from any sample of overt behavior to nervous nets is not unique, whereas, of imaginable nets, only one in fact exists, and may, at any moment, exhibit some unpredictable activity. Certainly for the psychiatrist it is more to the point that in such systems "Mind" no longer "goes more ghostly than a ghost." Instead, diseased mentality can be understood without loss of scope or rigor, in the scientific terms of neurophysiology. For neurology, the theory sharpens the distinction between nets necessary or merely sufficient for given activities, and so clarifies the relations of disturbed structure to disturbed function. In its own domain the difference between equivalent nets and nets equivalent in the narrow sense indicates the appropriate use and importance of temporal studies of nervous activity: and to mathematical biophysics the theory contributes a tool for rigorous symbolic treatment of known nets and an easy method of constructing hypothetical nets of required properties.

LITERATURE

Carnap, R. 1938. *The Logical Syntax of Language.* New York: Harcourt, Brace and Company.

Hilbert, D., und Ackermann, W. 1927. *Grundüge der Theoretischen Logik.* Berlin: J. Springer.

Russell, B., and Whitehead, A. N. 1925. *Principa Mathematica.* Cambridge: Cambridge University Press.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5. 1943

A STATISTICAL CONSEQUENCE OF THE LOGICAL CALCULUS OF NERVOUS NETS

H. D. LANDAHL, W. S. MCCULLOCH AND WALTER PITTS

THE UNIVERSITY OF CHICAGO, THE UNIVERSITY OF ILLINOIS, AND THE UNIVERSITY OF CHICAGO

A formal method is derived for converting logical relations among the actions of neurons in a net into statistical relations among the frequencies of their impulses.

Consider the neuron c_n upon which $c_1, \cdots, c_p$ have excitatory, and $c_{p+1}, \cdots, c_{p+q}$, inhibitory synapses. Let δ be the period of latent addition, so that c_n is excited at the time t if and only if the number of impulses along $c_1, \cdots, c_p$ which have concurred within an interval of duration δ about $t - 1$ exceeds θ_n, and none has occurred within δ along any of $c_{p+1}, \cdots, c_{p+q}$. Suppose that the sequences of impulses along $c_1, \cdots, c_p$ are statistically independent, and let ν_i be the mean frequency of impulses along c_i. Then the mean proportion of time that exactly r excitatory impulses arrive upon c_n within an interval of duration δ is given by

$$\sum_{i_r=i_{r-1}+1}^{i_r=p} \cdots \sum_{i_2=i_1+1}^{i_2=p} \sum_{i_1=1}^{i_1=p} \prod_{j=1}^{j=r} \delta^r \nu_{i_j} . \tag{1}$$

Consequently, the mean proportion of time Λ that the number of impulses concurring upon c_n within δ exceeds θ_n is obtained by summing (1) from $r = \theta_n$ to $r = p$. The intervals of duration δ within which impulses from any of $c_{p+1}, \cdots, c_{p+q}$ occur are intervals when c_n is inexcitable. The mean proportion of time when this is not the case is given by

$$\prod_{k=p+1}^{k=p+q} (1 - \delta \nu_k) . \tag{2}$$

The frequency ν_n is then the product of (2) by Λ, or

$$\nu_n = \delta^{-1} \prod_{k=p+1}^{k=p+q} (1 - \delta \nu_k) \sum_{r=\theta_n}^{r=p} \sum_{i_r=i_{r-1}+1}^{i_r=p} \cdots \sum_{i_2=i+1}^{i_2=p} \sum_{i_1=1}^{i_1=p} \prod_{j=1}^{j=r} \delta^r \nu_{i_j} . \tag{3}$$

If we compare equation (3) with expression (1) of the preceding

paper (McCulloch and Pitts, 1943), whose notation we shall use, and apply a straightforward inductive argument, we obtain the following

THEOREM.

Let $\mathcal{N}$ be a net of order zero, or, more generally, one for which

$$N_i(z) \equiv S_i[i = 1, \cdots, s]$$

is a solution wherein the S_i fulfill the conditions of Theorem X of the preceding paper. Let ν_j be the mean frequency of impulses in c_j, and let the expression A_i be generated out of S_i by the following rules:

(1) *Replace each* **N** *of S_i by '$\delta\,\nu$' with the same subscript.*

(2) *Replace every* **v** *by '+', every · by '×' and every* ∼ *by '1 −'.*

(3) *Replace the operators (z_i) and (Ez_i) respectively by $\prod_{z_1=0}^{z_2=z_1}$ and $\sum_{z_1=0}^{z_2=z_1}$ wherever they occur.*

(4) *Replace every occurrence of a predicate C_{mn} by a symbol for the function $f_{mn}(t)$ which is defined for all natural numbers t as unity when $t \equiv m \pmod n$, and otherwise zero.*

Then the frequencies of impulses in the c_i are given in terms of those of the peripheral afferents by the equations

$$\delta\,\nu_i = A_i[i = 1, \cdots, s]. \qquad (4)$$

The correspondence expressed by this theorem is exactly that of Boole between the algebra of logic and that of probability. It connects the logical calculus of the preceding paper with previous treatments of the activity of nervous nets in mathematical biophysics and with quantitatively measurable psychological phenomena. For these phenomena we can construct hypothetical nets by the powerful methods of the preceding calculus. The theorem then enables us to determine specific predictions from the quantitative characters of the stimulus to those of the response. These predictions can be compared with observations and, if necessary, the nets be altered until the consequent predictions are verified.

But this procedure leads to error whenever the assumptions leading to equation (3) are not fulfilled. This will be the case if the frequencies are too great, but the limit is many times the maximum observed. When the frequencies are too small, the statistical treatment, though valid, is of little help, but we may here conveniently use the logical calculus directly. The same obtains in microscopic physiologi-

cal analysis. The most important exceptions are those which arise from physiological factors that synchronize activities and thus restrict the domain of the validity of the assumption of statistical independences required in the derivation of equation (3).

LITERATURE

McCulloch, Warren S. and Walter Pitts. 1943. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bull. Math. Biophysics*, 5, 115-133.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

A MATHEMATICAL THEORY OF THE AFFECTIVE PSYCHOSES

JEROME Y. LETTVIN AND WALTER PITTS

BOSTON CITY HOSPITAL AND THE UNIVERSITY OF CHICAGO

The theory introduces two variables ϕ and ψ. The first represents the intensity of emotion, the second measures the intensity of activity. A set of integrodifferential equations is assumed to govern the variation of ϕ and ψ with respect to time. Since for increasing values of ϕ the conduct of the organism varies from great impassivity through a normal level of feeling to extremes of a circular depression or catatonic excitement; whereas an increase of ψ results in a transition from stupor to manic excitement, the solutions of the equations represent quantitative specifications of different psychotic states.

In the following discussion we shall develop a mathematical theory of the group of mental disorders which may be characterized in the following fashion: That the characteristic course of the disease may be essentially described in terms of the vicissitudes of two variables, the first representing the level of feeling, affect, or emotion in the organism, and the other the level of activity or conation. This group we consider to comprise the circular insanities, the reactive psychoses, and the catatonia of Kahlbaum. It may also be supposed to include affective disorders superimposed upon psychoses of another type, and perhaps also, with a more specific interpretation of the determining variables, certain forms of neurosis; but we shall not consider these latter cases in detail.

We shall denote the fundamental quantities of our theory, regarded as functions of the time t, by the symbols $\phi(t)$ and $\psi(t)$. For increasing values of $\phi(t)$ the conduct of the organism will vary from great impassivity to a normal level of feeling, then to normal strong emotion, and finally to the extremes of a circular depression or catatonic excitement. (We do not here take account of the quality of the emotion). If $\psi(t)$ rises indefinitely, the organism will pass from the greatest stupor and inactivity to normal conation, and ultimately to manic excitement or an idiomotor seizure. We shall measure these variables from an origin placed at the resting values normal for the organism, so that positive values of $\phi(t)$ and $\psi(t)$ correspond to supernormal affective and conative levels, and negative values to subnormal levels.

Consider the change $\Delta \phi$ of $\phi(t)$ during a small interval of time $t, t + \Delta t$. This quantity may be supposed to arise as the sum of three contributions, say $\Delta_1 \phi$, $\Delta_2 \phi$, $\Delta_3 \phi$. The first represents the influence of the contemporary environment at the time in question; for us this can be taken only as given, and specified as an empirical function $M(t)$. The second, $\Delta_2 \phi$, represents the effect of the previous experience of the organism; we may arrive at an expression for it in the following fashion.

The quantity $\Delta_2 \phi(t)$ for any subject is an average over a large number of terms, each of which represents the affectivity of the reaction he habitually makes to some particular kind of situation. Let us denote one of these terms by $\Delta_2 \phi_A(t)$, and consider the contribution to it of the events occurring during some very short interval of time in the past, say that from η until $\eta + d\eta$, where $0 \leqq \eta < t$. Denote this contribution by $\Delta_\eta \Delta_2 \phi_A(t)$.

We observe first that $\Delta_\eta \Delta_2 \phi_A(t)$ depends upon the degree to which the organism is concerned with situations of the given type at the time η; times when he is little or not at all concerned with them will have a negligible direct effect upon his future reactions to them, whereas periods of very great concern will usually have a very large effect. Denoting this concern by a function $Q_A(\eta)$, we shall embody it as a factor in the expression for $\Delta_\eta \Delta_2 \phi_A$, as is simplest to satisfy these conditions. Suppose then that $Q(\eta) > 0$, and let the subject's concern at time η with the given kind of situation have been attended with emotion, so that $\phi(\eta)$ is large. By the principles of learning, this occasion will cause recurrences of this type of matter to be treated with greater emotion than otherwise, or to manifest a higher value of ϕ; and this will be the larger, as $\phi(\eta)$ is greater. But if the subject's reactions at the time η have been of primarily conative type, with a high value for $\psi(\eta)$ and a correspondingly reduced emotion, the recurrences of the same kind of situation will also exhibit a smaller ϕ. This is to say, that $\Delta_\eta \Delta_2 \phi_A(t)$ increases with the value of $\phi(\eta)$ and decreases with $\psi(\eta)$, while the analogously defined $\Delta_\eta \Delta_2 \psi_A(t)$ behaves in an inverse fashion.

We shall suppose that this relation is a simple proportionality, or

$$\begin{aligned} \Delta_\eta \Delta_2 \phi_A(t) &= \beta' Q_A(\eta)[\phi(\eta) - \varepsilon_1 \psi(\eta)] \\ \Delta_\eta \Delta_2 \psi_A(t) &= \beta' Q_A(\eta)[\psi(\eta) - \varepsilon_2 \phi(\eta)] \end{aligned} \tag{1}$$

in which β', ε_1, and ε_2 are suitable constants of proportionality.

But the strength of any contribution to $\Delta_2 \phi_A(t)$ from the past will in general be less in an amount dependent upon its remoteness in time, by ordinary forgetting; so that we must multiply the left of

equation (1) by some decreasing function of the distance $t - \eta$ into the past. For this we shall select the function $e^{-\mu x}$, which seems to be a fair representation of the "law of forgetting", according to Ebbinghaus and others; the undetermined coefficient μ here allows us to reserve to psychoanalysts and other people who do not consider that an affecting experience ever becomes less efficient, the option of setting $\mu = 0$. We shall see, however, that this assumption requires greatly increased strength in the homeostatic mechanism of the organism to avoid continuously disturbed behavior.

The total magnitude of $\Delta_2 \phi(t)$ will be obtained by adding together all the quantities $e^{-\mu(t-\eta)} \Delta_\eta \Delta_2 \phi_A(t)$ corresponding to a set of past intervals of time which covers $(0, t)$ completely, and then summing quantities of this kind for each type of situation: Being linear they superpose, and if we allow the intervals into which we have divided a past to increase without limit in number, and their lengths to become uniformly indefinitely small, the error we have made in supposing that ϕ and ψ do not change in the interior of each interval approaches zero, and the sum will be transformed into an integral. We derive

$$\Delta_2 \phi(t) = \beta'' \int_0^t Q(\eta) e^{-\mu(t-\eta)} [\phi(\eta) - \varepsilon_1 \psi(\eta)] \, d\eta, \qquad (2)$$

together with a corresponding expression for $\Delta_2 \psi(t)$, where $Q(t)$ is the same kind of average of the Q_A's that $\phi(t)$ is of the ϕ_A's. This $Q(t)$ is in principle determinable empirically, in terms of the intuitive characterization we have given it above, but not easily so in fact. We may approximate it in somewhat the following fashion. On the average, those matters will tend to be of concern to a subject which excite his feelings, emotions, activity; and they will be of the more concern to him, the more they do so. This consideration will perhaps make it plausible to put $Q = \gamma \phi$ into equation (2) and its analogue for ψ, where γ is a constant of proportionality.

The last part of $\Delta \phi$, the term $\Delta_3 \phi(t)$, arises out of general homeostatic mechanisms of the organism. These tend to keep the values of ϕ and ψ close to the normal levels: whenever a deviation occurs from these levels, it brings into play restoring forces whose strength increases with the extent of the deviation and tends to reduce it. We may represent this component of $\Delta \phi$ and $\Delta \psi$ by a function f, so that $\Delta_3 \phi = - f(\phi)$, $\Delta_3 \psi = - f(\psi)$, where f is properly monotone and never negative. If we now pass to the limit as Δt becomes small, and set $\beta = \beta'$, we shall obtain the equations

$$\frac{d\phi}{dt} = \beta \int_0^t \phi(\eta) e^{-\mu(t-\eta)} [\phi(\eta) - \varepsilon_1 \psi(\eta)] \, d\eta + M(t) - f[\phi(t)],$$
$$\frac{d\psi}{dt} = \beta \int_0^t \phi(\eta) e^{-\mu(t-\eta)} [\psi(\eta) - \varepsilon_2 \phi(\eta)] \, d\eta + \bar{M}(t) - f[\phi(t)]. \tag{3}$$

These equations form the basis of our theory, and, with the addition of a few subsidiary hypotheses about organic interventions, everything is to be deduced from them. In the present study we shall in general treat M and $\bar{M}$, which represent the external fortunes or misfortunes of the organism, as if they consisted of a series of impulsive shocks between which they are relatively negligible. The "small pains and troubles of daily existence," in Schopenhauer's phrase, which most reasonably go into this term, may be taken account of in the determination of the normal values whence we measure ϕ and ψ.

Our first problem is to determine the simplest form for the restoring function $f(x)$ which will prevent the functions ϕ and ψ from increasing or decreasing without bound and assuming values which are physiologically meaningless. Considering $f(x)$ to be a polynomial, we see that its leading term must be an odd power of x, in order that the homeostasis may work against both positively and negatively abnormal values. If it is in addition not of the first degree, we may demonstrate its adequacy to limit ϕ and ψ in the following way.

Multiply the equations (3) by $e^{\mu t}$ throughout, differentiate, and cancel the exponential factor. We have then

$$\phi'' = -[\mu + f'(\phi)] \phi' + \beta \phi^2 - \beta \varepsilon_1 \phi \psi - \mu f(\phi)$$
$$\psi'' = -[\mu + f'(\psi)] \psi' + \beta \phi \psi - \beta \varepsilon_2 \phi^2 - \mu f(\psi), \tag{4}$$

with the initial conditions $\phi'(0) = -f[\phi(0)]$, $\psi'(0) = -f[\psi(0)]$, where we have dropped the terms in M and $\bar{M}$, since the variables cannot become infinite during the integrable impulses of which M and $\bar{M}$ consist, and at other times they obey the equations (4). Now time may be divided into two sorts of interval: those where $|\phi| \geqq |\psi|$, and those where the reverse inequality holds. Considering intervals of the former type, we find, upon rearranging the first of equations (4),

$$\phi' = \frac{\beta \phi^2 - \beta \varepsilon_1 \phi \psi}{\mu + f'(\phi)} - \frac{\phi''}{\mu + f'(\phi)} - \frac{\mu f(\phi)}{\mu + f'(\phi)}$$
$$= O(1) - O(\phi) - \frac{\phi''}{\mu + f'(\phi)}, \tag{5}$$

as $\phi \rightarrow \infty$ through intervals of the given type, under the given assumptions. Now we may distinguish these kinds of intervals of time. As ϕ increases through intervals of the first kind, we shall have $\phi'' = O[f'(\phi)]$; in this case, by (6) ϕ' is monotonically decreasing without bound for sufficiently large ϕ, and hence will ultimately become negative. As ϕ increases through intervals of the second kind, ϕ'' is negative and of a greater order than $f'(\phi)$: $\phi'' \leqq -\lambda f'(\phi)$, for some $\lambda > 0$. Since $f'(\phi)$ increases without bound with its argument, in intervals of this type ϕ' will also ultimately decrease monotonically and without bound, finally becoming negative. In intervals of the third kind, where $\phi'' \geqq \lambda f'(\phi)$, we discover the same from equation (5), so that we may conclude generally that $\phi(t)$ and $\psi(t)$ are bounded for all time in intervals where $|\phi| \geqq |\psi|$; by applying the same argument to the second equation (4), we establish the conclusion for all time without restriction. This of course does not allow us to exclude singularities in the solution, but since the right of (4) obviously satisfies the Lipshitz condition, that is no problem.

Among the forms of f which this result leaves open to us, we shall select the simplest, a cubic polynomial in which for the sake of symmetry the square term is omitted; we put

$$f(x) = \kappa_1 x + \kappa_3 x^3, \quad \kappa_1, \kappa_3 > 0.$$

We now discuss the character of the history determined by the equations 4. This may be represented by the motion of a particle in a plane, whose abscissa and ordinate at a time are respectively the values of ϕ and of ψ which hold at that time; the problem then becomes formally a dynamical system with two degrees of freedom to determine the path of such a particle. Since the coefficients of ϕ' and ψ' in equations (4) are always positive, the system is dissipative; since it is also bounded, a well-known theorem of dynamics (Birkhoff, 1927, Ch. I) allows us to conclude that it will asymptotically approach some stable point of equilibrium. The equilibria will be the real solution of the algebraic equations obtained by setting ϕ', ψ', ϕ'', ψ'' equal to zero in (4), which are

$$\begin{aligned} \beta\phi^2 - \beta\varepsilon_1\phi\psi &= \mu(\kappa_1\phi + \kappa_3\phi^3) \\ \beta\phi\psi - \beta\varepsilon_2\phi^2 &= \mu(\kappa_1\psi + \kappa_3\psi^3). \end{aligned} \tag{7}$$

We shall solve them under the assumption that $\varepsilon_1 = \varepsilon_2 = \varepsilon$: the motion does not change its character sharply if this is not precisely true, as an examination of the perturbation of first order will readily convince us; and in any case we expect these parameters to be of the

same order of magnitude. For numerical applications, a closer approximation can always be found.

The solutions are readily obtained, and, if we exclude those which are certainly imaginary, they are in general five in number:

1. A root $\phi_1 = \psi_1 = 0$, representing the normal values of the variables;
2. A pair of large equal roots
$$\phi_{21} = \frac{\beta(1-\varepsilon)}{2\mu\kappa_3} + \frac{1}{2\mu\kappa_3}[\beta^2(1-\varepsilon)^2 - 4\mu^2\kappa_1\kappa_3]^{\frac{1}{2}},$$
$$\psi_{21} = \phi_{21};$$
3. A second such pair:
$$\phi_{22} = \frac{\beta(1-\varepsilon)}{2\mu\kappa_3} - \frac{1}{2\mu\kappa_3}[\beta^2(1-\varepsilon)^2 - 4\mu^2\kappa_1\kappa_3]^{\frac{1}{2}},$$
$$\psi_{22} = \phi_{22};$$
4. One pair of large roots, negatives of one another:
$$\phi_{31} = \frac{\beta(1+\varepsilon)}{2\mu\kappa_3} + \frac{1}{2\mu\kappa_3}[\beta^2(1+\varepsilon)^2 - 4\mu^2\kappa_1\kappa_3]^{\frac{1}{2}},$$
$$\psi_{31} = -\phi_{31};$$
5. A smaller pair of such roots
$$\phi_{32} = \frac{\beta(1+\varepsilon)}{2\mu\kappa_3} - \frac{1}{2\mu\kappa_3}[\beta^2(1+\varepsilon)^2 - 4\mu^2\kappa_1\kappa_3]^{\frac{1}{2}},$$
$$\psi_{32} = -\phi_{32}.$$

Some or all of these roots, except of course ϕ_1, ψ_1, may happen to be imaginary, so that they do not in fact represent equilibria. This will be the case for the pairs (2) and (3) if the radical in them is negative, a condition which holds if and only if

$$1 - 2\sigma \leqq \varepsilon \leqq 1 + 2\sigma; \tag{8}$$

where $\sigma = \mu\sqrt{\kappa_1\kappa_3}/\beta$.

The equilibria corresponding to the roots (4) and (5) will exist under an inequality for the radical occurring there which may easily be transformed into

$$\begin{aligned} \varepsilon &> 2\sigma - 1 \\ \varepsilon &\leqq -\tfrac{1}{3} + \tfrac{2}{3}\sqrt{1 - 3\sigma^2}. \end{aligned} \tag{9}$$

For all sufficiently small σ, the inequalities (8) and (9) will obvi-

ously leave a range for ε in which all five equilibria exist. In perhaps the most common case, μ and κ_3 will be quite small, and β and κ_1, of the order of unity, so that σ is close to zero, the condition (9) is satisfied by virture of the sign of ε; and (8) requires only that ε should not be in the immediate neighborhood of unity. Still, if μ, κ_1, or κ_3 is quite large or β is small, both inequalities may fail, and all equilibria save the origin cease to exist; this latter is then necessarily stable. An organism whose parameters are so related cannot possibly exhibit disturbed behavior under any circumstances.

It is important to investigate the behavior of the solution in the neighborhood of the points of equilibrium, first, to determine their stability or instability, and second, to throw light on the fluctuations of the organism about the normal level which are insufficient to pass into permanent disturbance. If $\phi = \nu_1$, $\psi = \nu_2$ be a point of equilibrium, and if we set

$$\bar{\phi} = \phi - \nu_1, \quad \bar{\psi} = \psi - \nu_2,$$

we shall derive expressions for $\bar{\phi}$ and $\bar{\psi}$ of the form

$$\begin{aligned} \bar{\phi}(t) &= \sum_{i=1}^{4} \bar{A}_i e^{\lambda_i t} \\ \bar{\psi}(t) &= \sum_{i=1}^{4} \bar{A}_i e^{\lambda_i t} \end{aligned} \tag{10}$$

where the $\bar{A}_i$ are certain linear combinations of the A_i, these depend on the boundary conditions, and, of the system (4) for this equilibrium, the λ_i are the roots of the characteristic equation which is

$$\begin{vmatrix} \alpha_1 - \lambda & 0 & \gamma_1 & -\beta \varepsilon_1 \nu_1 \\ 0 & \alpha_2 - \lambda & \beta \nu_2 - 2\beta \varepsilon_2 \nu_1 & \gamma_2 \\ 1 & 0 & -\lambda & 0 \\ 0 & 1 & 0 & -\lambda \end{vmatrix} = 0$$

where

$$\begin{aligned} \alpha_1 &= -[\mu + f'(\nu_1)] \\ \alpha_2 &= -[\mu + f'(\nu_2)] \\ \gamma_1 &= 2\beta \nu_1 - \beta \varepsilon_1 \nu_2 - \mu f'(\nu_1) \\ \gamma_2 &= \beta \nu_1 - \mu f'(\nu_2), \end{aligned}$$

provided that all these roots are distinct and that none of them vanishes. This equation may be written in the form

$$\lambda^2(\lambda-\alpha_1)(\lambda-\alpha_2)-\lambda[\gamma_1(\lambda-\alpha_1)+\gamma_2(\lambda-\alpha_2)] + \gamma_1\gamma_2+\beta^2\varepsilon\nu_2(\nu_2-2\varepsilon\nu_1)=0\,. \tag{11}$$

Since $|\nu_1| = |\nu_2|$ for all the points of equilibrium, we have $\alpha_1 = \alpha_2 = \alpha$, and equation (11) is easily soluble, with roots

$$\lambda=\frac{\alpha}{2}\pm\tfrac{1}{2}[\alpha^2+2(\gamma_1+\gamma_2)\pm 2\sqrt{(\gamma_1+\gamma_2)^2-4\chi}]^{\frac{1}{2}}, \tag{12}$$

in which $\chi = \gamma_1\gamma_2 + \beta^2\varepsilon\nu_1(\nu_2 - 2\varepsilon\nu_1)$ and each of the four λ_i is obtained by a different choice for the two doubtful signs in expression (12). Now for the equilibria we are considering to be of stable type, in the sense that no sufficiently small deviation from that point will generate forces which tend to augment it, the condition is that none of the λ_i, or their real parts, if some of them be complex, shall exceed zero. This will be the case here if and only if

$$\alpha=-[\mu+f'(\nu)]<0\,,$$

which is always true,

$$\gamma_1+\gamma_2=3\beta\nu_1-2\mu\kappa_1-6\mu\kappa_3\nu^2_1<\beta\varepsilon\nu_2\,, \tag{13}$$

and

$$\chi=(3\mu\kappa_3\nu^2_1+\beta\varepsilon\nu_2+\mu\kappa_1-2\beta\nu_1)(3\mu\kappa_3\nu^2_2 + \mu\kappa_1-\beta\nu_2)+\beta^2\varepsilon\nu_1(\nu_2-2\nu_1\varepsilon)>0\,. \tag{14}$$

For the equilibrium at the origin, these reduce to

$$-2\mu\kappa_1<0\,,\quad \mu^2\kappa^2_1>0\,,$$

so that the origin is always a stable point, as we should expect. For the other equilibria, the conditions become extremely cumbersome when handled directly; we shall therefore approximate them in the following manner. The quantity $\sigma^2 = \mu^2\kappa_1\kappa_3/\beta^2$ may in general be expected to be quite small in comparison with unity; we may therefore expand in the expressions for the points of equilibrium in powers of σ^2, substitute in (13) and (14), and terminate the expansion at the first term that gives a determinate form to the conditions. We discover that for the points ϕ_{21}, ψ_{21} and ϕ_{31}, ψ_{31} we may neglect σ^2 itself, whereas for the smaller equilibria at ϕ_{22}, ψ_{22} and ϕ_{32}, ψ_{32} it is necessary to take account of terms in σ^2; and for one initial value, namely $\varepsilon = \frac{1}{2}$, we must proceed to terms in σ^4. The results are as follows: the condition (13) is satisfied

1) By ϕ_{21} , ψ_{21} unless $3/5 \leqq \varepsilon < 1$;
2) By ϕ_{22} , ψ_{22} unless $3 < \varepsilon \leqq 2 + \sqrt{10}$;
3) By ϕ_{31} , ψ_{31} if $\varepsilon \geqq 0$, i.e., always;
4) By ϕ_{32} , ψ_{32} if $\varepsilon \geqq 0$.

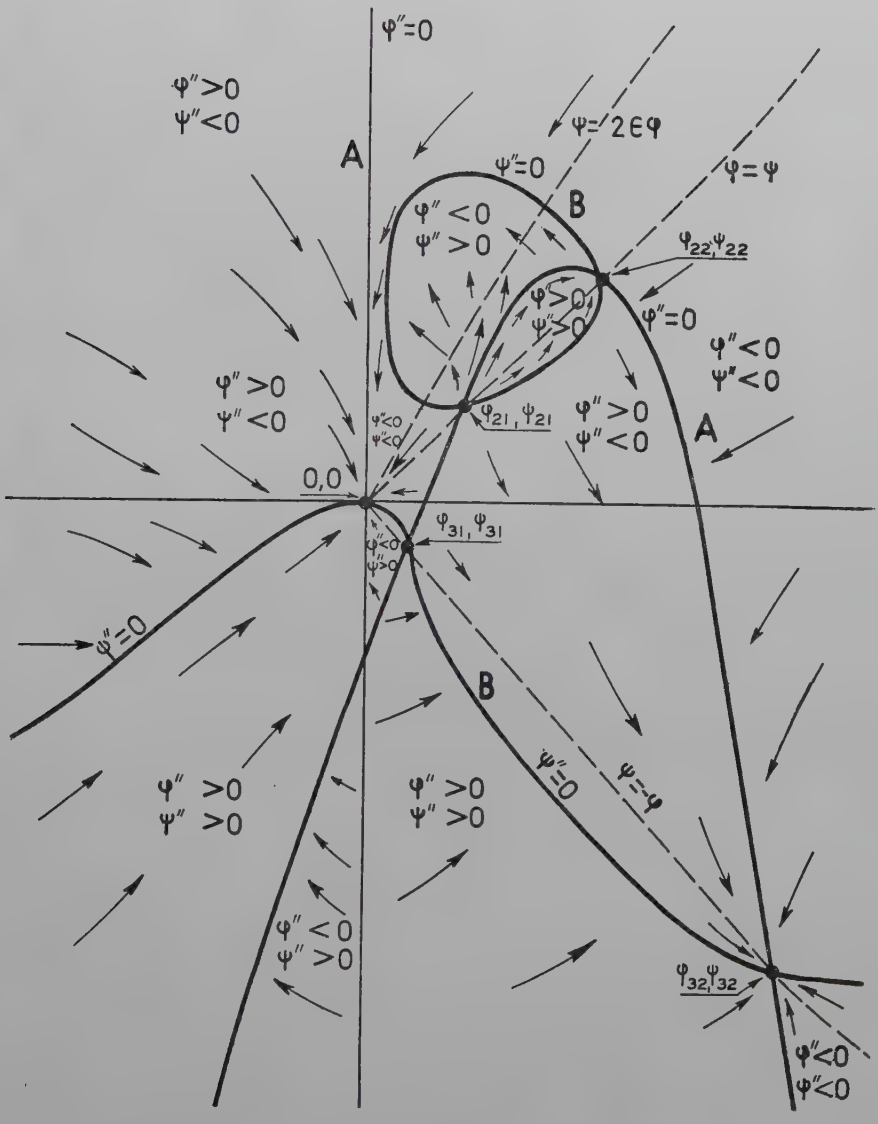

FIGURE 1

The inequality (14) is fulfilled

1) By ϕ_{21}, ψ_{21} unless $\frac{1}{2} < \varepsilon \leqq 1$;
2) By ϕ_{22}, ψ_{22} if $\varepsilon < \frac{1}{2}$;
3) By ϕ_{31}, ψ_{31} if $\varepsilon \geqq 0$;
4) By ϕ_{32}, ψ_{32} unless $\sqrt{7} - 2 < \varepsilon < 1$.

We may obtain a general insight into the qualitative character of the motion determined by our equations from the first figure. The degenerate algebraic curve consisting of a parabola A together with the axis of ψ divides the plane into the regions for which ϕ'' is positive and those for which it is negative, while the closed curve B together with the cubic parabola separate the regions of positive from those of negative ψ''. Clearly, the intersections of these curves will constitute the points of equilibrium. In each of these regions, we now have arrows, whose slope is the value of ψ''/ϕ'', which represent roughly the direction of motion of a particle placed there with no kinetic energy. The presence of kinetic energy and the dissipative forces will modify these considerations somewhat, and the initial conditions in (4) will superpose upon these forces an initial velocity with components $-f(\phi)$, $-f(\psi)$ toward the origin.

The difference in character between the stable and the unstable equilibria becomes very clear in the diagram. As remarked above, these are no periodic orbits in the large, and unless continually disturbed, the particle will ultimately settle toward one of the equilibria.

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

LITERATURE

Birkhoff, G. 1927. *Dynamical Systems.* New York: American Mathematical Society Colloquium Publication No. 9.

MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943
BULLETIN OF

ON SYNCHRONOUS SPORULATION WITH POSSIBLE REFERENCE TO MALARIAL PARASITES

ALSTON S. HOUSEHOLDER
THE UNIVERSITY OF CHICAGO

With the view of suggesting a possible mechanism, it is shown that sporulation may be synchronous, as with the malarial parasite, given the following condition: A metabolite which is essential to the growth of the parasite is periodically released from storage organs and dissipated from the blood stream of the host.

It is well known that the sporulation of the malarial plasmodium exhibits marked synchronism in virtually all species, and that while the period varies from species to species, it is in general some multiple of 24 hours, with exceptional instances of 12-hour multiples. Moreover, numerous experiments (Taliaferro and Taliaferro, 1934; Boyd, 1928, 1929 and 1933; Stauber, 1937) have shown that by altering the diurnal cycle of the host's behavior and environment in various ways, the synchronism of the sporulation can be disrupted and re-established at a new time in accordance with the new cycle imposed upon the host. On the other hand, while it is apparently rather easy to disrupt the initial synchronism, the re-establishment is more difficult, and there appear to be definite limits beyond which the period cannot be changed.

One can hardly question the conclusion, therefore, that the time of sporulation is somehow imposed upon the parasite by the host, but that the regulation is not complete, since there are quartan and tertian malarias as well as quotidian. There seems to be no generally accepted theory as to the mechanism involved, although a variety of suggestions have been made, based upon attempts to control independently this or that feature of the environment. But even if one should establish experimentally that, say, the time of sporulation could be influenced by the host's feeding habits, and not by his periods of activity and rest alone, one would still be faced with the task of interpreting this in terms of the physicochemical environment of the parasite, to say nothing of the difficulties that would be involved in conclusively establishing such a hypothesis in the first place.

We wish to consider the problem more directly from the point of view of the possible physiological mechanisms that might be capable

of exerting such a regulating effect. Many such mechanisms are possible *a priori*, but any one, considered quantitatively, must imply certain processes whose presence or absence can be established by direct physiological measurements. In this place we present one of these possibilities.

Rashevsky (1940) has shown that a unicellular organism may be expected to become unstable upon reaching a certain critical size, on account of the diffusion forces resulting from its own metabolic activity. It is indeed possible that, in some cases at least, other forces may contribute or may even anticipate the diffusion forces in bringing about the division, but the diffusion forces are shown to be adequate and to be effective at the right orders of magnitude. The critical radius determined by these forces depends upon the diffusion coefficients, the molecular weights and the metabolic rates for the various metabolites, as well as upon the surface tension and the temperature. Also the metabolic rates presumably depend, in part at least, upon the supply of the metabolite itself.

Moreover, the rate of growth of the organism must be dependent upon those same parameters, except possibly for surface tension, since growth can only occur through the intake of nutrient material from the surrounding medium. The simplest picture one can form of the entire process is that growth proceeds until the organism reaches a critical size, after which division occurs. The synchronism could then be due to factors of either or both of two types; factors which periodically lower the critical radius, or those which periodically yield a marked acceleration in the growth rate. We consider here the latter possibility, supposing the critical radius fixed.

The growth of the cell may be supposed to result from the excess of the intake of substance to be consumed over the elimination of waste products of the metabolism (Rashevsky, 1940), and the rate of intake is dependent, as remarked above, upon the supply. On the supposition that the rate of consumption of n metabolites is jointly proportional to the supply of these—their concentration within the cell—it has been further shown (Householder, 1943) that the various consumption rates are all approximately proportional to the available concentration of a single one of the metabolites, which thus acts as the "limiting factor." Then, if this limiting metabolite is one whose concentration in the blood stream varies sufficiently, the growth rate of the parasite, and hence its division time, will be largely controlled by this factor. The simplest supposition would be that the metabolite is released once during the day from storage organs within the host's body, after which it disappears exponentially. The time of release might have been established partly by conditioning from the host's

feeding habits, and, though stable for minor deviations in these, yet subject to variation as a result of thorough-going disruption.

Under these circumstances, if we take the concentration of the limiting metabolite to be proportional to e^{-at}, the relative increase in the size of the parasitic cell will be proportional to this, and its relative decrease through the process of elimination of wastes will be proportional to the rate of this elimination, which we suppose constant. Rashevsky's growth equation then has the form

$$2rdr/dt = a'e^{-at} - \beta r^2 , \tag{1}$$

where r represents the radius of the cell. The solution of this is

$$r^2 = r_0^2\, e^{-\beta t} - a(e^{-\beta t} - e^{-at}), \tag{2}$$

$$-a = a'/(\alpha - \beta), \quad \alpha \neq \beta , \tag{3}$$

or

$$r^2 = (r_0^2 + a't)\, e^{-at}, \quad \alpha = \beta . \tag{4}$$

It is enough for our purpose to discuss the case of $\alpha < \beta$. If $\alpha < \beta$, a as defined by (3) is positive, and by a simple change of scale we may take $a = 1$. Hence in this case we may write

$$r^2 = (r_0^2 - 1)\, e^{-\beta t} + e^{-at} . \tag{5}$$

It is natural, furthermore, to take 24 hours as the time unit. Hence equation (5) holds for $0 \leqq t \leqq 1$ if no division has occurred, after which the value r_0 must be replaced by a new value, the value of r^2 at $t = 1$. If this is now taken as a new origin of time, the equation holds again in the same form.

In writing equation (5), r_0 was taken as the value of r at the time $t = 0$, i.e., at the time of release of the metabolite postulated. As such, r_0^2 is essentially positive. However, division occurring at time t_c, an equation of the same form will hold for the schizont on the time-interval t_c, 1, provided r_0^2 is so chosen that the corresponding r at time t_c is equal to the radius of the schizont. The r_0^2 so determined will not necessarily be positive, but if not, this would only mean that any parasite, however small at time $t = 0$, would, by the time $t = t_c$, have grown to larger size.

For discussion it is convenient to introduce, in place of β, the parameter λ defined by

$$\beta = \alpha(1 + 1/\lambda). \tag{6}$$

Then in place of equation (5) we may write

$$r^2 = e^{-at}[1 - (1 - r_0^2)\, e^{-at/\lambda}] . \tag{7}$$

Also we have

$$\frac{d(r^2)}{dt} = \alpha e^{-\alpha t}\left[\frac{\lambda+1}{\lambda}(1-r_0^2)e^{-\alpha t/\lambda} - 1\right]. \qquad (8)$$

Now it is clear that at any time t, r^2 when regarded as a function of r_0^2 is monotonically increasing; moreover, the derivative, and hence the growth rate of the parasite, is monotonically decreasing. Hence if division does not intervene, two parasites which are initially unequal in size will remain so, and in the same sense, but the growth rate of the larger one is always less. In fact, unless

$$r_0^2 < (\lambda+1)^{-1}, \qquad (9)$$

growth will always be *negative*, i.e. the size decreases rather than increases. And even when relation (9) is satisfied, the right member of (8) always vanishes for a single value of $t > 0$, which may be less than unity, according to the value of r_0^2. Hence, in general, except for r_0^2 small and perhaps even negative, growth ultimately stops, and the size of the parasite then decreases until the next release of the metabolite. This decrease, however, is slow and need not be detectable.

If α/λ is large, say somewhat greater than unity, the second term within the brackets in equation (7) becomes negligible for $t = 1$. This means that the size of the parasite at the moment of the release of the metabolite is independent of r_0^2, or nearly so, and hence all parasites are then of the same size. This independence will be more or less complete according to the order of magnitude of α/λ, and the radius at this time is

$$r_1 = e^{-\alpha/2}.$$

Let us suppose then, as we must, that this is below the critical radius. Suppose, further, that

$$\frac{\alpha}{\lambda} > \log\frac{\lambda+1}{\lambda}.$$

Then the curve (7) for $r_0 = 0$ rises to a maximum for $t < 1$, and the same is true with any positive r_0^2 satisfying relation (9). If the square of the critical radius is less than the maximal r^2 given by (7) when $r_0^2 = e^{-\alpha}$, then division will always occur within 24 hours. Moreover, if

$$r_c \nu^{-1/3} < e^{-\alpha/2},$$

where ν is the number of schizonts and r_c the critical radius, then no second division can occur. However, if this is not the case, a second

division could occur. This would not be in just 12 hours, necessarily, but with a deviation of not more than an hour or so, a 12-hour period could easily be supposed. In either case the occurrence of synchronism is obvious.

If α/λ is small, there will be a spread in the radii at the time $t = 1$, but the division may still be synchronous. In fact, the following properties of the curves

$$r^2 = f(r_0^2, t) \tag{10}$$

are the only ones required, and these are easily secured with suitable values of the parameters α and λ:

1) $f(r_0^2, t)$ has a maximum in t for $t < 1$ for at least sufficiently large values of r_0^2;

2) for each value of t, $f(r_0^2, t)$ is monotonically increasing in r_0^2;

3) the condition

$$r_0^2 < f(r_0^2, 1) \tag{11}$$

is fulfilled for values of r_0 not too great.

Suppose, then, that the critical radius r_c is such that one of the curves (10) has r_c^2 as its maximum for $t = t'_c < 1$. Let this be defined by $r_0 = r'_0$ and let

$$r'^2_1 = f(r'^2_0, 1) > r'^2_0. \tag{12}$$

Then

$$\begin{aligned} r_c^2 &= f(r'^2_0, t'_c), \\ 0 &= f_t(r'^2_0, t'_c). \end{aligned} \tag{13}$$

Let r''_0 and r''_1 be defined by

$$\begin{aligned} r_c^2 \nu^{-2/3} &= f(r''^2_0, t'_c), \\ r''^2_1 &= f(r''^2_0, 1), \end{aligned} \tag{14}$$

and let

$$r_c^2 \nu^{-2/3} < r'^2_0, \quad r''^2_1 > r'^2_0. \tag{15}$$

Then no division can occur for $t > t'_c$, and if division occurs at $t = t'_c$, then r''^2_0 defines the curve of development of the schizonts, and r''_1 is their radius at $t = 1$. Conditions (15) insure, respectively, that at most one and at least one division occurs in one day, at any rate after the first day.

We can make this clear and establish the synchronism as follows. At the start of the second cycle every parasite has a radius r_0 satisfying $r''_1 \leqq r_0 \leqq r'_1$ so that every parasite divides during the second

cycle and in every one thereafter. Consider, then, any parasite of radius r_0 at the beginning of the second cycle, yielding schizonts of radius r_1 at the beginning of the next, r_2 at the next, and so on. Because of the monotonicity properties of the curves, if $r_1 < r_0$, then the second division occurs later than the first, hence $r_2 < r_1$ and the third occurs still later, and so sequentially. Likewise, if $r_1 > r_0$, the second division occurs earlier than the first, etc. Now if $r_0 = r'_1$, then necessarily $r_1 < r_0$, while if $r_0 = r''_1$, $r_1 > r_0$. Hence we can define two sequences, one decreasing with initial member $r_0 = r'_1$, one increasing with initial member $r_0 = r''_1$, and no member of the second can exceed any member of the first. Hence each sequence has a limit. It is easily seen that these limits are identical and define the steady-state value of the radius of the parasites at the beginning of the cycle.

We have supposed that each segmenter divides into ν exactly equal schizonts, and if this were the case, perfect synchronism would result in time. Actually, such perfect equality of division is not to be expected, and hence neither will perfect synchronism result. Smaller schizonts will be more tardy in dividing and some may fail entirely during the ensuing cycle. But with our hypothesis, all divisions which do occur must do so within a limited time interval, and the distribution of frequencies during this interval could be calculated from a knowledge of the size-frequencies among the schizonts. Finally, whereas we have considered explicitly only a quotidian malaria, it should be clear that the same mechanism but with different parameters ν, α and λ could yield equally well a tertian or a quartan period.

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Boyd, George H. 1928. "Induced Variations in the Asexual Cycle of *Plasmodium cathemerium*." *American Journal of Hygiene*, 9, 181-187.

Boyd, George H. 1929. "Experimental Modification of the Reproductive Activity of *Plasmodium cathemerium*." *Journal of Experimental Zoology*, 54, 111-126.

Boyd, George H. 1933. "Host Fatigue and Feeding in their Relation to the Reproductive Activity of *Plasmodium cathemerium* Hartman." *American Journal of Hygiene*, 18, 295-322.

Householder, Alston S. 1943. "Note On Rashevsky's Equation for Cellular Growth." *Bull. Math. Biophysics*, 5, 161-163.

Rashevsky, N. 1940. *Advances and Applications of Mathematical Biology*. Chicago: The University of Chicago Press.

Stauber, Leslie Alfred. 1937. *Factors Influencing Asexual Periodicity in Avian Malaria*. Chicago: University of Chicago Thesis.

Taliaferro, William H., and Lucy Graves Taliaferro. 1934. "Alteration in the Time of Sporulation of *Plasmodium brasilianum* in Monkeys by Reversal of Light and Dark." *American Journal of Hygiene*, 20, 50-59.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

A THEORY OF THE INDUCED SIZE EFFECT

ALSTON S. HOUSEHOLDER
THE UNIVERSITY OF CHICAGO

A quantitative formulation of Ogle's induced size effect is carried out on the assumption that vertical equality of the retinal images acts as a cue for localizing the subjective median plane as opposed to the horizontal retinal disparities which serve for localizing the subjective frontal plane; that this cue operates independently of such cues as are provided by ocular versions, and predominates when the versions are not great. The equation which embodies the theoretical predictions involves no parameters not directly measurable.

It is well known, and, from simple consideration of the horopter, easily predicted, that a cylindrical size lens placed with axis 90° before one eye induces a rotation of the subjective visual space away from this eye and toward the other. It is less well known, and not to be so predicted, that the lens placed with axis 180° induces a contrary rotation which is about equal in amount when the lens magnification is not too great. This phenomenon Ogle has designated the "induced size effect" and he has described it in a series of papers listed below. It is indicated in these papers that the effect must be somehow due to the disparity in the retinal images, but a quantitative formulation of this is not made. The purpose of the present paper is to make such a formulation and to show how, in at least the "pure" situations, Ogle's experimental results are easily predictable.

In general, Ogle's experimental procedure consists in setting up a movable plane at a fixed distance from the subject and asking the subject to adjust the plane in a position parallel to the frontal plane. Ogle notes (1940) that more consistent results are obtained if the fusion contours on the plane are "restricted to relatively small areas above or below the center of the plane." In practice these fusion contours appear to be small circles arranged in two horizontal lines above and below the center.

In the quantitative formulation contemplated, it is assumed that an object which forms retinal images that are equal in the vertical dimension are judged to lie on the *median plane,* whereas one which forms images that are unequal in the vertical dimension are judged to lie on the side of the median toward the eye having the larger image. Thus whereas the horizontal disparity of the retinal images

provides the primary cue for localizing objects with reference to the frontal plane, an entirely different cue, the vertical disparity, is utilized in localizing the median plane. It is assumed, therefore, that Ogle's subjects were actually locating their subjective median in his experiments on the induced size effect, and this seems to be borne out by Ogle's statement concerning the placing of the fusion contours not too far from the center of the movable plane which contained them. Certainly the two cues which in this case conflict, that provided by the horizontal and that provided by the vertical disparities, must both operate to some extent. But in placing the fusion contours not too far from the center of the plane, the horizontal disparities are minimized. Also the necessity for placing these above and below the center is plain, since otherwise no vertical disparity would be present.

To make the necessary calculations, we introduce a coordinate system with the x-axis joining the nodal points of the two eyes, the positive y-axis extending forward, and the positive z-axis upward in the medial plane. It is unnecessary to take account of the small distance between the nodal point and the center of rotation of the eye. If we introduce as the unit of length half the interocular distance, then the left and right eyes have the coordinates $(-1,0,0)$ and $(+1,0,0)$. An object whose extremities are at $(x,y,0)$ and (x,y,z) subtends at the two eyes angles λ_1 and λ_2 given by

$$\tan \lambda_i = z\,[(x \pm 1)^2 + y^2]^{-\frac{1}{2}}, \tag{1}$$

and if a size lens of magnification μ is placed before the left eye, the angle λ'_1 is given by

$$\tan \lambda'_1 = \mu z\,[(x+1)^2 + y^2]^{-\frac{1}{2}}. \tag{2}$$

These angles are equal and the object is judged to be in the median plane provided

$$\mu\,[(x-1)^2 + y^2]^{\frac{1}{2}} = [(x+1)^2 + y^2]^{\frac{1}{2}}, \tag{3}$$

or provided

$$\left(x - \frac{\mu^2+1}{\mu^2-1}\right)^2 + y^2 = \left(\frac{2\mu}{\mu^2-1}\right)^2. \tag{4}$$

Thus as we vary the true frontal plane containing the fixation point, the intersection with this of the subjective median describes a circle whose center is collinear with the two eyes. The intercepts of this circle with the x-axis are at

$$(\mu \pm 1)/(\mu \mp 1).$$

Evidently for

$$y > 2\mu/(\mu^2 - 1)$$

image equality becomes impossible and the cue fails.

Experimentally the magnifications used are of only a few per cent. If we set

$$\mu = 1 + u\,, \tag{5}$$

so that $100u$ is the percentage magnification, (4) becomes, on neglecting u^2,

$$\left(x - \frac{1+u}{u}\right)^2 + y^2 = \left(\frac{1+u}{u}\right)^2, \tag{6}$$

the approximating circles all passing through the origin. For a magnification of 10%, the error in the radius as calculated this way is only 3%. The breakdown mentioned above occurs with the subjective median at an angle of 45°, with

$$x = y = (1+u)/u\,.$$

If we set

$$\lambda = \frac{u}{2(1+u)}, \tag{7}$$

then for a given y,

$$\frac{x}{y} = \lambda y\,[1 + \lambda^2 y^2 + 2\lambda^4 y^4 + \cdots]\,, \tag{8}$$

and this is the tangent of the angle of deviation of the subjective from the true median at the fixation distance y. For small angles we may take this as the radian measure of the angle itself.

We may conclude, therefore, that on the basis of our assumptions, the angle of rotation of the median plane for near fixation is approximately proportional to the percentage magnification, to the fixation distance, and to the interocular distance, this angle being kept less than 45°. At any given fixation distance, however, if the magnification is too great, the cue breaks down, since the retinal images cannot thus be equalized. We find, in fact, from Ogle's data that the curves graphing angle against magnification, for a given fixation distance, actually are nearly linear up to about 6% or so, after which they show a decline. The dependence upon the fixation distance is borne out also by curves obtained for 20 cm. and for 75 cm. distance. It is especially to be noted that the theory contains no undetermined parameters but gives absolute measurements.

The first of the accompanying figures shows the circles (6) drawn for magnification up to 10%. The second shows Ogle's data for two subjects (Ogle, 1938), both right and left eyes, compared with the theoretical curve. The fixation distance is 40 cm., and this was

assumed to be 6 times the interocular distance: $y = 12$. Actually the pupillary distances were given as 62 mm. and 61 mm. in the two cases, making y slightly greater than this. These data are not carried to the "breakdown" point, and in general this occurs earlier than is predicted. In this connection, however, we may note, first, that such an angle already substantially exceeds the "small angle" limits for the optical formulas employed in the calculation of the magnification, and second, that additional cues such as ocular versions, not accounted for in the formula, are bound to become increasingly sensible as the angle increases.

This leads up to the general problem of how the various cues combine to yield the geometric structure of the perceived visual space. Among the purely geometrical cues, disregarding such things as aerial perspective, there are at least five: the ocular versions, the vertical equality or inequality of the retinal images (on the present theory), convergence, the horizontal disparity, and, presumably, accommodation. The first two are concerned with the location of the medial plane, the second being more sensitive for smaller deviations, but only for these. The third and possibly the last serve to localize the general region focussed with reference to the observer. The fourth gives depth to this region and gives the cues for small differences. Under "normal" conditions, all of these would cooperate to yield a more or less accurate representation of the objective space under examination. Under "abnormal" conditions, whether the abnormality is physiological (ametropia, aniseikonia, strabismus) or is artificially produced by instrumentation, the cues may be conflicting and the resulting impression will be dependent upon the influence of the abnormality upon each of the several cues and upon the relative degrees of dominance of the various cues. We are here assuming the complete dominance of one of these for small disparities. The present formulation being accepted, it now becomes possible to employ simple algebra to calculate the influence of each of the five cues mentioned when taken alone, and thus by artificially varying these among themselves, it is possible to determine experimentally the degree to which each contributes to the total spatial picture in any given situation. Thus, for example, the geometrical and the induced size effects when placed in opposition by a spherical size lens appear to combine by simple addition so that they just neutralize one another when the magnification is not too great. On the other hand, the cues provided by the ocular versions do not "take hold" to overcome the conflicting induced effect until a minimal critical angle is reached. After this the version cue becomes increasingly dominant and overcomes that provided by the retinal inequality.

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

LITERATURE

Ogle, Kenneth N. 1938. "Induced Size Effect. I. A New Phenomenon in Binocular Space Perception Associated with the Relative Sizes of the Images of the Two Eyes." *Archives of Opthalmology*, **20**, 604-623.

Ogle, Kenneth N. 1939. "Induced Size Effect. II. An Experimental Study of the Phenomenon with Restricted Fusion Stimuli." *Archives of Opthalmology*, **21**, 604-625.

Ogle, Kenneth N. 1939. "Induced Size Effect. III. A Study of the Phenomenon as Influenced by Horizontal Disparity of the Fusion Contours." *Archives of Opthalmology*, **22**, 613-635.

Ogle, Kenneth N. 1940. "Induced Size Effect with the Eyes in Asymmetric Convergence." *Archives of Opthalmology*, **23**, 1023-1038.

Ogle, Kenneth N. 1940. "The Induced Size Effect." *Journal of the Optical Society of America*, **30**, 145-151.

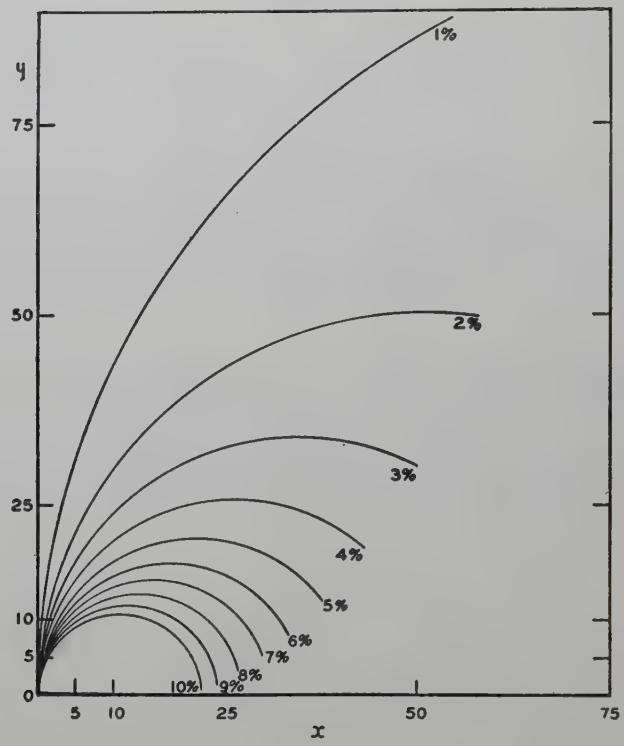

FIGURE 1

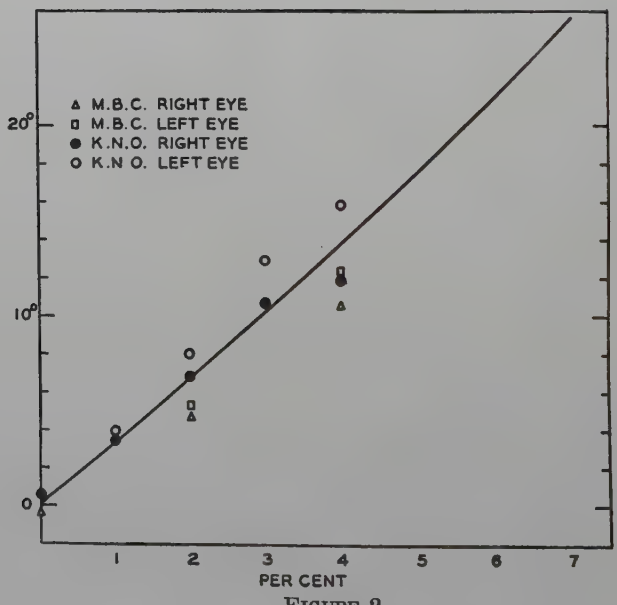

FIGURE 2

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

NOTE ON RASHEVSKY'S EQUATION FOR CELLULAR GROWTH

ALSTON S. HOUSEHOLDER
THE UNIVERSITY OF CHICAGO

The equation for growth by intussusception involves the product of the internal concentrations of n metabolites. This product is shown, in general, to satisfy a certain algebraic equation of the n-th degree. Approximate solutions are exhibited for a somewhat wider class of cases than are considered by Rashevsky.

Rashevsky (1940) derives an equation for the rate of growth of a cell by intussusception based upon the assumption that the rate at which each metabolite is transformed into the cell substance is jointly proportional to the concentrations of all the metabolites within the cell. With slow growth, the metabolic rate being regarded as always at the steady state, previous derivations give the (average) internal concentrations of the various metabolites in terms of their external concentrations, the diffusion coefficients and permeabilities, and the radius of the cell. Thus there are available the $2n$ equations

$$-q_l = k_l \, \Pi \, \bar{c} \, , \tag{1}$$

$$\bar{c}_l = c_{0l} + \Lambda_l \, q_l \, , \tag{2}$$

for the determination of the internal concentrations $\bar{c}_l$ and the consumption rates $-q_l$. The symbol $\Pi \, \bar{c}$ designates the product of all the $\bar{c}$'s, and the coefficients Λ_l are functions of the parameters mentioned.

In deriving the growth equation, Rashevsky evades the solution of these equations by resorting to an approximation. It is the purpose of this note to discuss this solution, which requires the solution of a single algebraic equation of the n-th degree, and to consider other possible approximations. To obtain the single equation we set

$$\begin{aligned} \alpha_l \, \Lambda_l \, k_l \, \Pi \, c_0 &= c_{0l} \, , \\ \Pi \, (\bar{c}/c_0) &= x \, , \end{aligned} \tag{3}$$

and obtain

$$q_l = -k_l \, x \, \Pi \, c_0 \, , \tag{4}$$

$$\bar{c}_l / c_{0l} = 1 - x/\alpha_l \, . \tag{5}$$

If we multiply together the n equations (5) we have then

$$x = \Pi(1 - x/\alpha). \tag{6}$$

This is the algebraic equation referred to. The $\bar{c}_l$ enter the growth equation only through the product of all of them, so that it is this quantity x alone which is needed.

It is no restriction to suppose that the n substances are so enumerated that

$$\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n .$$

Then from (5) it is clear that

$$x \leqq \alpha_1 .$$

If we graph the line and the curve

$$y = x , \qquad y = \Pi(1 - x/\alpha), \tag{7}$$

these intersect at the point whose abscissa and ordinate are equal to the required value of x. The curve has its y-intercept at 1 and the x-intercepts at α_l, so that the required intersection lies on the arc of the curve between $(0, 1)$ and $(\alpha_1, 0)$. Consideration of concavity indicates that the intersection lies between the secant joining these two points and the tangent at $(0, 1)$. The intersections of these lines with the line (7) are easily found and yield the limits

$$1/(1 + n/\alpha) \leqq x \leqq 1/(1 + 1/\alpha_1), \tag{8}$$

where α is the harmonic mean of the α_l. These inequalities become equalities in the special case when α_1 is so small that the other α's do not contribute appreciably to the harmonic mean. From (3) it is evident that this case of a small α_1 occurs when c_{01} is small or Λ_1 is large in comparison with k_l. In this case, that is, the one substance alone becomes a "limiting factor" because of its limited external concentration (small c_{01}), or its slow diffusion (large Λ_1) in comparison with the k_1 which determines its consumption rate.

Rashevsky's approximation was based essentially upon the assumption that the α_l be, all of them, nearly equal and small. In this case, x is small and equal to the common value of the α_l. This approximation and the one here suggested are exclusive and nearly complementary. All metabolites with large α's can evidently be neglected, and if the remaining ones have α's which are small and nearly equal, we return to Rashevsky's case for these which remain. If the small α's are comparable but still fairly different, however, although the degree of the equation may be reduced, other methods are required for a convenient approximation.

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

LITERATURE

Rashevsky, N. 1940. *Advances and Applications of Mathematical Biology.* Chicago: The University of Chicago Press.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

ON THE ORIGIN OF LIFE

N. RASHEVSKY
THE UNIVERSITY OF CHICAGO

On the basis of the recently proposed new fundamental equation of mathematical biophysics, a suggestion is made for a theory of the formation of a primitive cell from nonliving material. The discussion includes a suggestion for a quantitative formulation of the degree of biological organization. It is shown that according to the fundamental equation of mathematical biophysics, organization of the nonliving material may spontaneously increase under certain conditions, leading to a formation of a primitive organism. This process however is a very slow one, requiring time intervals of several years or even decades. This may account for the failure in observing or artificially producing spontaneous generation.

We shall apply here the fundamental equation of mathematical biophysics (Rashevsky, 1943c) to the problem of formation of a primitive cell from nonliving material. Using the same notations as before, we have

$$f_i \frac{dx_i}{dt} = \tau \frac{\partial F}{\partial x_i} - \frac{\partial H}{\partial x_i}. \tag{1}$$

One of the fundamental characteristics of any organism is its distinctness from the environment. In the nonliving liquid material the natural state is that of a uniform distribution of different solutes in a solvent. If a cell or any other organism is to be formed out of such a mixture, some of the molecules of this mixture must be segregated spatially and arranged into definite geometric patterns which enable us to perceive that arrangement as a unit distinct from the quasi-uniform environment. We may speak of this process as organization.

A great deal of loose talking and writing has been done about organization; but, to our knowledge, no one has yet attempted a simple *quantitative* description or definition of it. It is possible that no general useful definition can be given at present; in that case, a special definition in this particular instance is indicated, at least as a temporary expedient. This is what we propose to do.

Consider first a purely hypothetical abstract case: namely that a cell consists of a unit, composed of certain substances that may in general freely spread through the surrounding solution but that ac-

tually remain confined to a geometrically clearly limited region of space. We may define the degree of organization in this case as follows:

Let $2R_0$ be the linear dimensions of the nonliving medium, for simplicity assumed to be roughly spherical in shape. Let the r_0 be the radius of the cell, which is also assumed to be spherical. Let

$$R_0 >> r_0 . \tag{2}$$

When the constituent material of the cell is spread uniformly throughout the volume $\frac{4}{3}\pi R_0^3$ of the environment, the cell as such is non-existent. For this case we may set the organization as being zero. If on the other hand, all of the constituent material of the cell is concentrated inside a sphere of radius r_0, and none outside, we shall put the organization of that state as equal to one. Intermediate values of organization will be found when some of the concentration of the material inside a sphere of radius r_0 is greater than outside, but when that outside concentration is not zero.

Denote by x the average concentration of the constituent material in the cell, by y—its average concentration outside. Considering the constant amount of the total material in the system, we have, because of inequality (2), denoting by C a constant:

$$xr_0^3 + yR_0^3 = C . \tag{3}$$

From the above it follows that when $x = y$, the organization O of the system is zero. When $y = 0$, then $O = 1$. We may conveniently define O by:

$$O = \frac{x - y}{x} . \tag{4}$$

Equation (4) then determines the degree of organization for any intermediate stage.

Thus we illustrate on this simple example how the degree of organization can be defined quantitatively.

In other, more complex cases, different definitions of O may have to be given. For instance, we may consider a cell as a sack consisting of a membrane composed of a substance m, which holds inside the other constituents of the cell, preventing them from spreading out. Denoting by x_m and y_m the concentrations of the substance m in the membrane of thickness δ, that is, in the volume $4\pi r_0^2\delta$, and in the surrounding medium correspondingly, and by x_c and y_c the average concentrations of the other cell constituents, we may define O by:

$$O = \tfrac{1}{2}\frac{(x_c - y_c)^2}{x_c^2} + \tfrac{1}{2}\frac{(x_m - y_m)^2}{x_m^2}. \qquad (5)$$

The organization is again zero for $x_c = y_c$; $x_m = y_m$, that is, for a uniform mixture. It is equal to one for a completely formed cell. The quadratic, rather than linear, terms are used in order to avoid the absurd situation in which a nonuniform distribution of material would give $O = 0$. Instead of equation (4) we might also put

$$O = \left(\frac{x - y}{x}\right)^2, \qquad (6)$$

in order to avoid negative value of O. Inasmuch as all the above expressions are used only as illustrations, we shall confine ourselves here to expression (4).

The increase of organization is accompanied by an increase in free energy, although the total energy may not change. Changes in total energies, connected with changes in organization are in most cases rather small. Therefore for simplicity we shall neglect them here. But the change in energy transformation may be very large when the organization changes. We do not gain much energy by crushing a cell, but we either completely stop or strongly reduce many energy producing or consuming metabolic reactions. We may therefore consider, at least theoretically, the case in which H does not depend on O, but F does. Reasons for that dependence may be various. Reaction rates may depend in many ways on the concentrations; or the confinement of catalists to a small volume may increase their efficiency much more rapidly than the increase of their concentration. In general we have $F = F(O)$. *As an illustration only,* we shall consider here the case

$$F = K O, \qquad (7)$$

K being a constant.

If F is an increasing function of O, then considering O as the coordinate describing the configuration of the system, we see from equation (1) that dO/dt will be positive, and that O thus will increase.

In order to apply equation (1) to our problem we now have to compute the viscous force $f(dx_i/dt)$.

This force will be due to the diffusion resistance and may be estimated in the following manner.

If the average transport velocity of the molecules is v, then, remembering that the *average* radius of the sphere through which

the substance is flowing when converging toward the forming cell is $\frac{1}{2} R_0$, we have

$$\pi R_0^2 y v = \frac{4}{3} \pi R_0^3 \frac{dy}{dt}$$

or

$$v = \frac{4R_0}{3y} \frac{dy}{dt}. \tag{8}$$

Considering each molecule as a sphere with an *average* radius a, the force of resistance per molecule is $6\pi\eta a v$, where η is the viscosity of the solvent. Denoting by M the average molecular weight of the molecule, and by N Avogadro's number, the number of molecules per cm^3 becomes yN/M, and therefore the force of resistance per cm^3 is

$$\frac{8\pi\eta a N R_0}{M} \frac{dy}{dt}. \tag{9}$$

Hence, because of inequality (2), the total resistance force in the system is equal to

$$\frac{32\pi^2\eta a N R_0^4}{3M} \frac{dy}{dt}. \tag{10}$$

Putting

$$A = \frac{C}{r_0^3}; \quad B = \frac{R_0^3}{r_0^3} \; (>> 1); \quad P = \frac{32\pi^2\eta a N}{3M}; \tag{11}$$

we find from equation (3) and (4) that

$$O = \frac{A - (B+1)y}{A - By}. \tag{12}$$

Solving equation (12) for y, differentiating with respect to time and introducing the result into expression (10) we find for the total force of viscous resistance the expression

$$f \frac{dO}{dt} = \frac{APR_0^4}{(1 + B - BO)^2} \frac{dO}{dt}. \tag{13}$$

Introducing expression (7) and (13) into equation (1) we now find

$$\frac{PR_0^4 A}{(1 + B - BO)^2} \frac{dO}{dt} = \tau K. \tag{14}$$

Putting

$$\frac{APR_0^4}{B} = \alpha, \tag{15}$$

and integrating equation (14) with the initial condition $O = 0$ for $t = 0$, we find

$$O = \frac{(1+B)\tau Kt}{B\tau Kt + \dfrac{\alpha BB}{1+B}}. \tag{16}$$

The time t_1 at which O becomes equal to 1 is given by

$$t_1 = \frac{\alpha B}{\tau K(1+B)}. \tag{17}$$

If, as may be expected, τ is very small (Rashevsky, 1943c), then t_1 will be very large. Thus the process of formation of a cell may require several years or decades. This may account for the failure in observing or artificially producing spontaneous generation.

LITERATURE

Rashevsky, N. 1943a. "Outline of a New Mathematical Approach to General Biology: I." *Bull. Math. Biophysics*, 5, 33-47.

Rashevsky, N. 1943b. "Outline of a New Mathematical Approach to General Biology: II." *Bull. Math. Biophysics*, 5, 49-64.

Rashevsky, N. 1943c. "Note on the Hamiltonian Principle in Biology and in Physics." *Bull. Math. Biophysics*, 5, 65-68.

INDEX TO VOLUME 5

INDEX OF AUTHORS

INDEX OF SUBJECTS

BOOK REVIEW

SCOPE OF THE BULLETIN

1. The Bulletin is devoted to publications of research in Mathematical Biophysics, as contributing to the physicomathematical foundations of biology in their most general scope.

2. Papers published in the Bulletin cover physicomathematical theories as well as any other mathematical treatments of biological phenomena, with the exception of purely statistical studies.

3. Mathematical studies in physics or in borderline fields in which a direct connection with biological problems is pointed out are also accepted.

4. Emphasis is put upon the mathematical developments, but a description and discussion of experimental work falls also within the scope of the Bulletin provided that description or discussion is made in close connection with mathematical developments *contained in the same paper.*

5. Outside of the scope of the journal are papers of purely statistical nature or papers concerned only with empirical equations.

PREPARATION OF MANUSCRIPTS

All manuscripts should be typewritten double space. Equations should be numbered on the right consecutively. *Do not* use such notations as "equation 2a" or, "equation 5′ " etc. References should be all given at the end of the paper, arranged in alphabetic order by authors and for each author chronologically, following *exactly* the style used in the Bulletin. In text, reference should be made by giving in parentheses the name of the author followed by the year of publication. In case of several publications by the same author in the same year, use notations "1940a", "1940b", etc.

In writing equations, slanted lines should be used wherever possible.

Every paper is to be preceded by a short abstract.

Drawings should be prepared in a professional manner on white paper or tracing cloth following as closely as possible the style used in the Bulletin. They should be drawn approximately twice the scale of the finished reproduction. Lettering should be made by means of a lettering guide. If desired by the authors, drawings may be made at the Editorial Office according to author's instructions, at cost.